GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 00:14:43 ; Search time 373.587 Seconds
(without alignments)
10012.591 Million cell updates/sec

Title: US-09-001-737-7
Perfect score: 1661
Sequence: 1 GAATTCGGCTTCATATGGCA.....TGGGCGGATAAGCCGAATTC 1661

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1661	100.0	1661	20	AAX86155	DNA encoding a Str
	2	1622.2	97.7	1635	24	ABN69160	Streptococcus poly
	3	1413.4	85.1	2683	22	AAH56861	S. pyogenes groEL
	4	1140.4	68.7	1306	22	AAH56862	Antibiotic resista
	5	1122.6	67.6	1305	22	AAH56865	Antibiotic resista
c	6	1063.4	64.0	2155561	24	ABN71527	Streptococcus poly
	7	1062.4	64.0	1620	24	ABN69159	Streptococcus poly
	8	1052	63.3	1654	20	AAX86153	DNA encoding a Str
	9	1039.2	62.6	1623	22	AAH90800	CFE 104 coding seq
	10	1038.4	62.5	1926	22	AAF25036	Nucleotide sequenc
	11	1037	62.4	2107	22	AAH56860	Antibiotic resista
	12	1036	62.4	1647	22	AAH90906	2CFE 104 coding se
	13	1035.4	62.3	5365	19	AAV52210	Streptococcus pneu
	14	1030.6	62.0	2401	22	AAH56859	S. pneumoniae groE
	15	955	57.5	2365589	24	ABA90521	Genomic sequence o
	16	942.6	56.7	3625	20	AAX12979	Enterococcus faeca
	17	832	50.1	1901	24	ABQ70705	Listeria monocytog
	18	824	49.6	213251	24	ABQ67193	Listeria innocua c
	19	786.4	47.3	1957	22	AAH56868	S. aureus RN4220 g
	20	773	46.5	1958	22	AAH56867	Methicillin resist
	21	760.6	45.8	1570	24	ABQ69121	Listeria monocytog
c	22	756.4	45.5	3885	22	AAH54519	S. epidermidis gen
	23	755.4	45.5	1620	22	AAH52698	S. epidermidis ope
	24	735.6	44.3	1623	24	ABN92405	Staphylococcus epi
	25	722.4	43.5	2712	22	AAH56866	S. aureus groE ope
	26	704.2	42.4	1647	18	AAT69201	Lawsonia intracell
	27	704.2	42.4	4972	18	AAT69203	Lawsonia intracell
	28	654.2	39.4	96109	22	AAF28548	Genomic fragment #
	29	615.8	37.1	640681	24	ABA92787	Buchnera sp. genom
	30	608.2	36.6	613	22	AAH56864	Antibiotic resista
	31	604.2	36.4	609	22	AAH56863	Antibiotic resista
	32	590.6	35.6	1665	20	AAX86152	DNA encoding a Str
	33	589.8	35.5	1647	21	AAA48498	E. coli groEL codi
	34	589.8	35.5	1647	22	AAC90805	Escherichia coli G
	35	589.8	35.5	1647	24	ABL96324	Nucleotide sequenc
	36	589.8	35.5	1741	22	AAH56858	E. coli K12 strain
	37	589.8	35.5	2155	24	AAD33423	E. coli groESL gen
	38	589.8	35.5	4524	21	AAA48500	Escherichia coli g
	39	583.4	35.1	2032	22	AAH56856	Multidrug resistan
	40	577.2	34.8	2006	22	AAH56857	Uropathogen E. col
	41	568.2	34.2	1662	20	AAX86154	DNA encoding a Str
	42	552.8	33.3	1635	21	AAZ54509	Neisseria meningit
c	43	551.6	33.2	910715	20	AAX20248	Borrelia burgdorfe
	44	550	33.1	2465	12	AAQ13136	Hyp operon. Chlam
	45	536.8	32.3	1635	21	AAA13000	Neisseria meningit

ALIGNMENTS

RESULT 1
AAX86155
ID AAX86155 standard; DNA; 1661 BP.
XX
AC AAX86155;
XX
DT 22-SEP-1999 (first entry)
XX
DE DNA encoding a Streptococcus pyogenes heat shock protein (Hsp)60-2.
XX
KW Heat shock protein; Hsp60-2; immune response; immunological carrier;
KW cancer control; tumour; sarcoma; cancer; gene therapy; ss.
XX
OS Streptococcus pyogenes.
XX
PN WO9935270-A1.
XX
PD 15-JUL-1999.
XX
PF 29-DEC-1998; 98WO-CA01203.
XX
PR 31-DEC-1997; 97US-0001737.
XX
PA (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI Mizzen L, Wisniewski J;
XX
DR WPI; 1999-430397/36.
DR P-PSDB; AAY23904.
XX
PT New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT useful in vaccines, as carriers for other immunogens, as anticancer

PT agents and for diagnosis
XX
PS Claim 3; Fig 4A-B; 176pp; English.
XX
CC The present sequence encodes a heat shock protein, designated Hsp60-2.
CC The protein, its fragments, variants and fusion proteins, are
CC used to elicit or enhance an immune response against Streptococcus,
CC and to elicit a similar response to a target antigen fused to the
CC protein. Unlike other immunological carriers, Hsp60 proteins are not
CC immunosuppressive so provide an increased response to any conjugated or
CC fused antigen. Also, where used for cancer control, they lack the side
CC effects associated with endotoxins. They can also be used to detect
CC specific antibodies and in treatment or prevention of tumours
CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or
CC liver). The Hsp60 polynucleotide is used for recombinant production
CC of the protein, as a source of primers and probes for detecting
CC streptococci in standard hybridization/amplification assays, and
CC therapeutically in gene therapy vectors.
XX
SQ Sequence 1661 BP; 512 A; 308 C; 407 G; 434 T; 0 other;

Query Match 100.0%; Score 1661; DB 20; Length 1661;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GAATTCGGCTTCATATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GAATTCGGCTTCATATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGG 60

Qy      61 TGCGCGGAGTTGATATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGCGCGGAGTTGATATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCA 120

Qy     121 ATGTTGTTCTTGAAAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATGTTGTTCTTGAAAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTG 180

Qy     181 CTAAAGAGATCGAATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTAAAGAGATCGAATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAG 240

Qy     241 TGGCTTCTAAAACCAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TGGCTTCTAAAACCAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACAC 300

Qy     301 AAGCCATTGTTCATGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AAGCCATTGTTCATGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCC 360

Qy     361 GTCGAGGCATTGAAACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTCGAGGCATTGAAACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAAC 420

Qy     421 CTGTATCTGGCAAGGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTGTATCTGGCAAGGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAG 480

Qy     481 TTGGAGAGTATATCTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TTGGAGAGTATATCTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCG 540

Qy     541 AAGAATCTCGAGGTATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAGAATCTCGAGGTATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTG 600

Qy     601 GTTACCTGTCTCAATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTTACCTGTCTCAATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACC 660

Qy     661 CATTTATCTTAATCACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CATTTATCTTAATCACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTG 720

Qy     721 AGGAAGTTCTTAAAACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AGGAAGTTCTTAAAACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAG 780

Qy     781 CACTTCCAACCCTTGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CACTTCCAACCCTTGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAG 840

Qy     841 CGCCAGGATTTGGTGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CGCCAGGATTTGGTGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTG 900

Qy     901 GTACAGTGATTACAGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GTACAGTGATTACAGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTG 960

Qy     961 GACAGGCTGCTAAGATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GACAGGCTGCTAAGATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAA 1020

Qy    1021 GTTCAGAAGCTATTGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GTTCAGAAGCTATTGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTT 1080

Qy    1081 CTGACTTTGACCGTGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CTGACTTTGACCGTGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTG 1140

Qy    1141 TTATCAAAGTAGGAGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTATCAAAGTAGGAGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGG 1200

Qy    1201 ATGCTCTAAATGCTACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 ATGCTCTAAATGCTACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAG 1260

Qy    1261 CACTTATTACGGTTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CACTTATTACGGTTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTG 1320

Qy    1321 GACGTAACATTGTGCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GACGTAACATTGTGCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTG 1380

Qy    1381 GGTACGAAGGCTCCGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GGTACGAAGGCTCCGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTA 1440

Qy    1441 ATGCTGCAACAGGTGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 ATGCTGCAACAGGTGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAG 1500

Qy    1501 TAACACGATCAGCGCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 TAACACGATCAGCGCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAG 1560

Qy    1561 CAGTTGTTGCTAATAAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 CAGTTGTTGCTAATAAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGG 1620

Qy    1621 ATCCAGGAATGATGGGTGGGATGGGCGGATAAGCCGAATTC 1661
          |||||||||||||||||||||||||||||||||||||||||
Db    1621 ATCCAGGAATGATGGGTGGGATGGGCGGATAAGCCGAATTC 1661
```

RESULT 2
ABN69160
ID ABN69160 standard; DNA; 1635 BP.
XX
AC ABN69160;
XX
DT 01-JUL-2002 (first entry)
XX

DE Streptococcus polynucleotide SEQ ID NO 6233.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR P-PSDB; ABP28529.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 7; Page 3785-3786; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (Sl), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyrogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 1635 BP; 504 A; 303 C; 397 G; 431 T; 0 other;

Query Match 97.7%; Score 1622.2; DB 24; Length 1635;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      13 ATATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTG 72
           ||||||||||||| || |||||||||||||||||||||||||||||||||||||||||||
Db       5 ATATGGCAAAAGATATTAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTG 64

Qy      73 ATATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTG 132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 ATATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTG 124

Qy     133 AAAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCG 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125 AAAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCG 184

Qy     193 AATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAA 252
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185 AATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAA 244

Qy     253 CCAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTC 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     245 CCAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTC 304

Qy     313 ATGAAGGACTAAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTG 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     305 ATGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTG 364

Qy     373 AAACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCA 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     365 AAACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCA 424

Qy     433 AGGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATA 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     425 AGGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATA 484

Qy     493 TCTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAG 552
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     485 TCTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAG 544

Qy     553 GTATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTC 612
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     545 GTATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTC 604

Qy     613 AATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAA 672
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     605 AATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAA 664

Qy     673 TCACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTA 732
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     665 TCACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTA 724

Qy     733 AAACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCC 792
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     725 AAACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCC 784

Qy     793 TTGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTG 852
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     785 TTGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTG 844

Qy     853 GTGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTA 912
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     845 GTGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTA 904

Qy     913 CAGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTA 972
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     905 CAGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTA 964

Qy     973 AGATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTA 1032
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     965 AGATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTA 1024

Qy    1033 TTGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACC 1092
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1025 TTGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACC 1084

Qy    1093 GTGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAG 1152
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1085 GTGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAG 1144

Qy    1153 GAGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATG 1212
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1145 GAGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATG 1204

Qy    1213 CTACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGG 1272
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1205 CTACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGG 1264

Qy    1273 TTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTG 1332
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1265 TTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTG 1324

Qy    1333 TGCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCT 1392
```

```
Db    1325 TGCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCT 1384

Qy    1393 CCGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAG 1452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1385 CCGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAG 1444

Qy    1453 GTGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAG 1512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1445 GTGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAG 1504

Qy    1513 CGCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTA 1572
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1505 CGCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTA 1564

Qy    1573 ATAAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGA 1632
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1565 ATAAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGA 1624

Qy    1633 TGGGTGG 1639
           |||| ||
Db    1625 TGGGCGG 1631
```

RESULT 3
AAH56861
ID AAH56861 standard; DNA; 2683 BP.
XX
AC AAH56861;
XX
DT 06-SEP-2001 (first entry)
XX
DE S. pyogenes groEL gene partial sequence SEQ ID NO:7.
XX
KW Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;
KW microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;
KW Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;
KW antibacterial; antiviral; antiproliferative; antisense therapy;
KW microbial infection; ds.
XX
OS Streptococcus pyogenes.
XX
PN WO200136625-A2.
XX
PD 25-MAY-2001.
XX
PF 20-NOV-2000; 2000WO-CA01347.
XX
PR 18-NOV-1999; 99US-0166249.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Dugourd D;
XX
DR WPI; 2001-355633/37.
XX
PT Novel antisense compounds targeting nucleic acid encoding groEL or
PT groES gene of microorganism, which hybridize with and inhibit
PT expression of the genes, useful to inhibit growth of microorganism
PT having the genes -
XX
PS Disclosure; Fig 7; 110pp; English.
XX
CC The present invention specifically claims AAH56368 to AAH56832 which are
CC antisense oligonucleotides to nucleotide sequences encoding groE. More
CC generally, antisense compounds (I) comprising antisense oligonucleotides
CC of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat
CC shock protein (HSP)60) (GL) and groES (HSP10) (GS) gene from a
CC microorganism, where the antisense compound is complementary to GL or
CC GS of a microorganism and specifically hybridises with and inhibits the
CC expression of GL or GS, is claimed. (I) have antibacterial, antiviral
CC and antiproliferative activities, and can be used in antisense therapy
CC and for inhibition of expression of groES or groEL. (I) are useful for
CC inhibiting expression of GL or GS in cells or tissues in vitro. (I) are
CC also useful for inhibiting the growth of a microorganism, or inhibiting
CC the expression of GL or GS gene in a microorganism (a bacterial cell or
CC a virus) having a GL or GS gene which involves administering to the
CC microorganism or to a cell infected with the microorganism, (I). (I) are
CC also useful for treating a mammalian pathological condition mediated by
CC the microorganisms which involves identifying a eukaryotic organism
CC having a pathological condition mediated by microorganisms having a GL
CC or GS gene and administering (I) such that the growth of microorganism
CC is inhibited. The antisense compounds are utilised for diagnostics,
CC therapeutics, prophylaxis and as research reagents and kits, e.g., to
CC prevent or delay microbial infections in humans. They are also useful as
CC molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854
CC represent PCR primers for groE sequences which are used in the
CC exemplification of the present invention. AAH56855 to AAH56870 represent
CC groE nucleotide sequence given in the present invention.
XX
SQ Sequence 2683 BP; 808 A; 490 C; 593 G; 792 T; 0 other;

Query Match 85.1%; Score 1413.4; DB 22; Length 2683;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1423; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy     201 GATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAACCAATGAT 260
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GATCATTTCGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAACCAATGAT 60

Qy     261 ATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCATGAAGGA 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCATGAAGGA 120

Qy     321 CTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGAAACAGCA 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGAAACAGCA 180

Qy     381 ACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAAGGAAGCT 440
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     181 ACAGCAACAGCTGTTGAAGCCTTGAAAGCCGTTGCTCAACCTGTATCTGGCAAGGAAGCT 240

Qy     441 ATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATATCTCAGAA 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATATCTCAGAA 300

Qy     501 GCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGTATGGAA 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGTATGGAA 360

Qy     561 ACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAATACATG 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAATACATG 420

Qy     621 GTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATCACGGAT 680
           |||||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||
Db     421 GTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATTATGGAT 480

Qy     681 AAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAAACCAAC 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAAACCAAC 540

Qy     741 CGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCTTGTCTTG 800
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     541 CGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCCCTTCCAACCCTTGTCTTG 600

Qy     801 AACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGTGATCGT 860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGTGATCGT 660

Qy     861 CGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACAGAGGAT 920
           |||||||||||||||||||||||||||||||||| |||||||| |||||||||||||||||
Db     661 CGTAAAGCTATGCTTGAAGACATTGCTATCTTGACTGGTGGTACTGTGATTACAGAGGAT 720

Qy     921 CTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAGATTACA 980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     721 CTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAGATTACA 780

Qy     981 GTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATTGCTAAC 1040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATTGCTAAC 840

Qy    1041 CGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGTGAAAAA 1100
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CGCATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGTGAAAAA 900

Qy    1101 CTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGAGCTCCA 1160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGAGCTCCA 960

Qy    1161 ACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCTACACGT 1220
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     961 ACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCCACACGT 1020

Qy    1221 GCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTTATTGAA 1280
           |||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db    1021 GCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTGATCGAA 1080

Qy    1281 AAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGTGCTTCGT 1340
           |||||||| |||||||||||||||||||||| |||||||||||||||||||||||||||
Db    1081 AAAGTAGCTGCTCTTGAGCTTGAGGGCGATGACGCTACTGGACGTAACATTGTGCTTCGT 1140

Qy    1341 GCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCCGTAGTT 1400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCCGTAGTT 1200

Qy    1401 ATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGTGAGTGG 1460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 ATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGTGAGTGG 1260

Qy    1461 GTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCGCTTCAA 1520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCGCTTCAA 1320

Qy    1521 AATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAATAAACCT 1580
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    1321 AATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAATAAGCCT 1380

Qy    1581 GAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGATGGGTGG 1639
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||| ||
Db    1381 GAACCAGCTGCGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGATGGGCGG 1439


RESULT 4
AAH56862
ID   AAH56862 standard; DNA; 1306 BP.
XX
AC   AAH56862;
XX
DT   06-SEP-2001  (first entry)
XX
DE   Antibiotic resistant S. pyogenes groEL gene partial sequence SEQ ID:8.
XX
KW   Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;
KW   microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;
KW   Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;
KW   antibacterial; antiviral; antiproliferative; antisense therapy;
KW   microbial infection; ds.
XX
OS   Streptococcus pyogenes.
XX
PN   WO200136625-A2.
XX
PD   25-MAY-2001.
XX
PF   20-NOV-2000; 2000WO-CA01347.
XX
PR   18-NOV-1999;   99US-0166249.
XX
PA   (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI   Wright JA,  Young AH,  Dugourd D;
XX
DR   WPI; 2001-355633/37.
XX
PT   Novel antisense compounds targeting nucleic acid encoding groEL or
PT   groES gene of microorganism, which hybridize with and inhibit
PT   expression of the genes, useful to inhibit growth of microorganism
PT   having the genes -
XX
PS   Disclosure; Fig 8; 110pp; English.
XX
CC   The present invention specifically claims AAH56368 to AAH56832 which are
CC   antisense oligonucleotides to nucleotide sequences encoding groE. More
CC   generally, antisense compounds (I) comprising antisense oligonucleotides
CC   of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat
CC   shock protein (HSP)60) (GL) and groES (HSP10) (GS) gene from a
CC   microorganism, where the antisense compound is complementary to GL or
CC   GS of a microorganism and specifically hybridises with and inhibits the
CC   expression of GL or GS, is claimed. (I) have antibacterial, antiviral
CC   and antiproliferative activities, and can be used in antisense therapy
CC   and for inhibition of expression of groES or groEL. (I) are useful for
CC   inhibiting expression of GL or GS in cells or tissues in vitro. (I) are
CC   also useful for inhibiting the growth of a microorganism, or inhibiting
CC   the expression of GL or GS gene in a microorganism (a bacterial cell or
CC   a virus) having a GL or GS gene which involves administering to the
CC   microorganism or to a cell infected with the microorganism, (I). (I) are
CC   also useful for treating a mammalian pathological condition mediated by
CC   the microorganisms which involves identifying a eukaryotic organism
CC   having a pathological condition mediated by microorganisms having a GL
CC   or GS gene and administering (I) such that the growth of microorganism
CC   is inhibited. The antisense compounds are utilised for diagnostics,
CC   therapeutics, prophylaxis and as research reagents and kits, e.g., to
CC   prevent or delay microbial infections in humans. They are also useful as
CC   molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854
CC   represent PCR primers for groE sequences which are used in the
CC   exemplification of the present invention. AAH56855 to AAH56870 represent
CC   groE nucleotide sequence given in the present invention.
XX
SQ   Sequence 1306 BP; 407 A; 235 C; 308 G; 356 T; 0 other;

  Query Match            68.7%;  Score 1140.4;  DB 22;  Length 1306;
  Best Local Similarity  99.9%;  Pred. No. 7.6e-289;
  Matches 1141;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     498 GAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGTATG 557
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 GAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGTATG 61

Qy     558 GAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAATAC 617
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 GAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAATAC 121

Qy     618 ATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATCACG 677
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 ATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATCACG 181

Qy     678 GATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAAACC 737
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 GATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAAACC 241

Qy     738 AACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCTTGTC 797
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 AACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCTTGTC 301

Qy     798 TTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGTGAT 857
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 TTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGTGAT 361

Qy     858 CGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACAGAG 917
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACAGAG 421
```

```
Qy    918 GATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAGATT 977
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 GATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAGATT 481

Qy    978 ACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATTGCT 1037
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 ACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATTGCT 541

Qy   1038 AACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGTGAA 1097
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 AACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGTGAA 601

Qy   1098 AAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGAGCT 1157
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602 AAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGAGCT 661

Qy   1158 CCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCTACA 1217
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 CCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCTACA 721

Qy   1218 CGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTTATT 1277
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    722 CGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTTATT 781

Qy   1278 GAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGTGCTT 1337
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    782 GAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGTGCTT 841

Qy   1338 CGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCCGTA 1397
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    842 CGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCCGTA 901

Qy   1398 GTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGTGAG 1457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    902 GTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGTGAG 961

Qy   1458 TGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCGCTT 1517
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    962 TGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCGCTT 1021

Qy   1518 CAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAATAAA 1577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1022 CAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAATAAA 1081

Qy   1578 CCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGATGGGT 1637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1082 CCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGATGGGC 1141

Qy   1638 GG 1639
          ||
Db   1142 GG 1143
```

```
RESULT 5
AAH56865
ID   AAH56865 standard; DNA; 1305 BP.
XX
AC   AAH56865;
XX
DT   06-SEP-2001  (first entry)
XX
DE   Antibiotic resistant S. pyogenes groEL gene partial sequence SEQ ID:11.
XX
KW   Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;
KW   microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;
KW   Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;
KW   antibacterial; antiviral; antiproliferative; antisense therapy;
KW   microbial infection; ds.
XX
OS   Streptococcus pyogenes.
XX
PN   WO200136625-A2.
XX
PD   25-MAY-2001.
XX
PF   20-NOV-2000; 2000WO-CA01347.
XX
PR   18-NOV-1999;  99US-0166249.
XX
PA   (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI   Wright JA,  Young AH,  Dugourd D;
XX
DR   WPI; 2001-355633/37.
XX
PT   Novel antisense compounds targeting nucleic acid encoding groEL or
PT   groES gene of microorganism, which hybridize with and inhibit
PT   expression of the genes, useful to inhibit growth of microorganism
PT   having the genes -
XX
PS   Disclosure; Fig 11; 110pp; English.
XX
CC   The present invention specifically claims AAH56368 to AAH56832 which are
CC   antisense oligonucleotides to nucleotide sequences encoding groE. More
CC   generally, antisense compounds (I) comprising antisense oligonucleotides
CC   of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat
CC   shock protein (HSP)60) (GL) and groES (HSP10) (GS) gene from a
CC   microorganism, where the antisense compound is complementary to GL or
CC   GS of a microorganism and specifically hybridises with and inhibits the
CC   expression of GL or GS, is claimed. (I) have antibacterial, antiviral
CC   and antiproliferative activities, and can be used in antisense therapy
CC   and for inhibition of expression of groES or groEL. (I) are useful for
CC   inhibiting expression of GL or GS in cells or tissues in vitro. (I) are
CC   also useful for inhibiting the growth of a microorganism, or inhibiting
CC   the expression of GL or GS gene in a microorganism (a bacterial cell or
CC   a virus) having a GL or GS gene which involves administering to the
CC   microorganism or to a cell infected with the microorganism, (I). (I) are
CC   also useful for treating a mammalian pathological condition mediated by
CC   the microorganisms which involves identifying a eukaryotic organism
CC   having a pathological condition mediated by microorganisms having a GL
CC   or GS gene and administering (I) such that the growth of microorganism
CC   is inhibited. The antisense compounds are utilised for diagnostics,
CC   therapeutics, prophylaxis and as research reagents and kits, e.g., to
CC   prevent or delay microbial infections in humans. They are also useful as
CC   molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854
CC   represent PCR primers for groE sequences which are used in the
CC   exemplification of the present invention. AAH56855 to AAH56870 represent
CC   groE nucleotide sequence given in the present invention.
XX
SQ   Sequence 1305 BP; 406 A; 238 C; 312 G; 349 T; 0 other;

  Query Match            67.6%;  Score 1122.6;  DB 22;  Length 1305;
  Best Local Similarity  99.2%;  Pred. No. 3.6e-284;
  Matches 1128;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    503 TATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGTATGGAAAC 562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGTATGGAAAC 60

Qy    563 AGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAATACATGGT 622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAATACATGGT 120

Qy    623 CACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATCACGGATAA 682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATCACGGATAA 180

Qy    683 AAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAAACCAACCG 742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAAACCAACCG 240

Qy    743 TCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCTTGTCTTGAA 802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCCCTTCCAACCCTTGTCTTGAA 300
```

```
Qy     803 CAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGTGATCGTCG 862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     301 CAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGTGATCGCCG 360

Qy     863 TAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACAGAGGATCT 922
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACAGAGGATCT 420

Qy     923 AGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAGATTACAGT 982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAGATTACAGT 480

Qy     983 TGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATTGCTAACCG 1042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATTGCTAACCG 540

Qy    1043 TATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGTGAAAAACT 1102
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     541 TATTGCACTGATTAAATCGCAATTAGAGACAACAACTTCTGACTTTGACCGTGAAAAACT 600

Qy    1103 ACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGAGCTCCAAC 1162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGAGCTCCAAC 660

Qy    1163 AGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCTACACGTGC 1222
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     661 AGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGACGCTCTAAATGCTACACGTGC 720

Qy    1223 AGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTTATTGAAAA 1282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTTATTGAAAA 780

Qy    1283 AGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGTGCTTCGTGC 1342
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     781 AGTAGCAGCTCTTGAACTTGAGGGCGATGATGCTACTGGACGTAACATTGTGCTTCGTGC 840

Qy    1343 TCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCCGTAGTTAT 1402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCCGTAGTTAT 900

Qy    1403 TGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGTGAGTGGGT 1462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGTGAGTGGGT 960

Qy    1463 TGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCGCTTCAAAA 1522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCGCTTCAAAA 1020

Qy    1523 TGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAATAAACCTGA 1582
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||| |||||
Db    1021 TGCAGCTTCTGTAGCTAGTCTGATTTTGACAACAGAAGCAGTTGTTGCTAATAAGCCTGA 1080

Qy    1583 ACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGATGGGTGG 1639
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    1081 ACCAGCTGCGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGATGGGCGG 1137


RESULT 6
ABN71527/c
ID   ABN71527 standard; DNA; 2155561 BP.
XX
AC   ABN71527;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Streptococcus polynucleotide SEQ ID NO 10967.
XX
KW   Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW   group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW   antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS   Streptococcus sp.
XX
PN   WO200234771-A2.
XX
PD   02-MAY-2002.
XX
PF   29-OCT-2001; 2001WO-GB04789.
XX
PR   27-OCT-2000; 2000GB-0026333.
PR   24-NOV-2000; 2000GB-0028727.
PR   07-MAR-2001; 2001GB-0005640.
XX
PA   (CHIR-) CHIRON SPA.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI   Tettelin H;
XX
DR   WPI; 2002-352536/38.
XX
PT   New Streptococcus protein for the treatment or prevention of infection
PT   or disease caused by Streptococcus bacteria, such as meningitis, and
PT   for detecting a compound that binds to the protein -
XX
PS   Claim 8; Page 4196-4488; 4525pp; English.
XX
CC   The invention relates to a protein (ABP25413-ABP30895) from group B
CC   streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC   (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC   the specification. The proteins have antibacterial and antiinflammatory
CC   activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC   antibodies that bind (I) are used in the manufacture of medicaments for
CC   the treatment or prevention of infection or disease caused by
CC   Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC   Nucleic acids encoding (I) are used to detect Streptococcus in a
CC   biological sample. (I) is used to determine whether a compound binds to
CC   (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC   used as a vaccine or diagnostic composition. The disease caused by
CC   Streptococcus that is prevented or treated may be meningitis. Nucleic
CC   acid encoding (I) may be used to recombinantly produce (I) and may be
CC   used in gene therapy. Antibodies to (I) are used for affinity
CC   chromatography, immunoassays, and distinguishing/identifying
CC   Streptococcus proteins.
XX
SQ   Sequence 2155561 BP; 695741 A; 385320 C; 382958 G; 691542 T; 0 other;

  Query Match           64.0%;  Score 1063.4;  DB 24;  Length 2155561;
  Best Local Similarity  79.2%;  Pred. No. 2.3e-267;
  Matches 1275;  Conservative    0;  Mismatches  331;  Indels    3;  Gaps    1;

Qy      14 TATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTGA 73
           |||||||||||| || ||||||||||||||||| ||  | || |||||||| || ||||||
Db 2024270 TATGGCAAAAGATATTAAATTTTCAGCAGATGCCCGCTCAGCAATGGTGCGTGGTGTTGA 2024211

Qy      74 TATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTGA 133
           ||| ||||| ||||| ||||||||||| | ||||||||||| || ||||||||||||||||
Db 2024210 TATTTTAGCTGATACAGTCAAAGTAACATTAGGTCCTAAAGGCCGTAATGTTGTTCTTGA 2024151

Qy     134 AAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCGA 193
           |||||| |||||||||||| |||||||| ||||| || || || || |||||||||| || ||
Db 2024150 AAAAGCATTTGGTTCTCCTTTAATTACAAATGATGGTGTGACAATTGCTAAAGAAATTGA 2024091

Qy     194 ATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAAC 253
           |||||||||| |||||||| |||||||| ||| | ||||| |||||||||||| |||||
Db 2024090 GCTAGAAGATCACTTTGAAAATATGGGAGCTAAACTTGTGTCAGAAGTGGCTTCAAAAAC 2024031

Qy     254 CAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCA 313
           |||||||||| || ||||| || || ||||| || |||||||| ||||| ||||| |
Db 2024030 TAATGATATTGCAGGGGATGGCACTACAACTGCTACTGTTTTGACCCAAGCTATTGTACG 2023971

Qy     314 TGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGA 373
           ||||| || |||||||| || ||||| || ||||| ||||| || ||||| || ||||||
Db 2023970 GGAAGGTCTTAAAAATGTAACTGCAGGGGCAAATCCGATTGGCATTCGTCGTGGTATTGA 2023911

Qy     374 AACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAA 433
```

```
Db 2023910 AACAGCTGTTTCAGCAGCAGTTGAAGAGCTAAAAGAGATTGCACAACCAGTTTCAGGCAA 2023851

Qy    434 GGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATAT 493

Db 2023850 AGAAGCTATTGCTCAAGTTGCGGCTGTGTCTTCACGTTCTGAAAAAGTTGGGGAATACAT 2023791

Qy    494 CTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGG 553

Db 2023790 TTCTGAAGCTATGGAGCGCGTGGGTAATGATGGTGTTATCACTATTGAAGAATCGCGAGG 2023731

Qy    554 TATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCA 613

Db 2023730 TATGGAAACAGAGCTTGAAGTTGTGGAAGGAATGCAGTTTGACCGTGGGTACTTGTCACA 2023671

Qy    614 ATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAAT 673

Db 2023670 GTATATGGTAACTGATAACGAGAAAATGGTCTCTGAACTTGAGAATCCGTATATCCTTAT 2023611

Qy    674 CACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAA 733

Db 2023610 TACAGATAAGAAAATTTCAAATATCCAAGAAATTTTACCATTATTAGAAGAGGTTCTTAA 2023551

Qy    734 AACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCT 793

Db 2023550 AACAAATCGTCCGTTGCTAATCATCGCTGATGATGTTGATGGAGAAGCTCTCCCAACGCT 2023491

Qy    794 TGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGG 853

Db 2023490 TGTTCTTAACAAAATTCGTGGAACTTTCAATGTCGTAGCTGTTAAAGCGCCTGGATTTGG 2023431

Qy    854 TGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTAC 913

Db 2023430 TGATCGTCGTAAAGCCATGCTGGAAGATATTGCTATCCTAACAGGAGGAACTGTCGTTAC 2023371

Qy    914 AGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAA 973

Db 2023370 TGAAGACCTTGGTTTAGACTTAAAAGATGCTACTATGCAAGTTTTAGGACAGTCTGCTAA 2023311

Qy    974 GATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTAT 1033

Db 2023310 AGTAACAGTAGATAAAGATTCTACTGTTATTGTCGAAGGTGCCGGTGACTCATCAGCAAT 2023251

Qy   1034 TGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCG 1093

Db 2023250 TGCTAATCGCGTAGCTATCATTAAGTCACAGATGGAGGCTACAACTTCTGATTTTGATCG 2023191

Qy   1094 TGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGG 1153

Db 2023190 TGAAAAATTACAAGAACGACTTGCTAAGTTAGCCGGTGGTGTAGCAGTAATTAAAGTTGG 2023131

Qy   1154 AGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGC 1213

Db 2023130 TGCAGCGACTGAAACAGAATTAAAAGAGATGAAACTTCGCATCGAAGATGCGTTAAATGC 2023071

Qy   1214 TACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGT 1273

Db 2023070 AACGCGTGCTGCAGTTGAAGAAGGTATTGTTTCAGGTGGAGGTACGGCTCTTGTGAACGT 2023011

Qy   1274 TATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGT 1333

Db 2023010 TATTGAAAAAGTAGCGGCACTGAAACTTAATGGTGATGAGGAGACTGGACGTAATATTGT 2022951

Qy   1334 GCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTC 1393

Db 2022950 TCTTCGTGCTCTCGAAGAGCCTGTTCGTCAAATTGCTTACAATGCTGGATATGAAGGTTC 2022891

Qy   1394 CGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGG 1453

Db 2022890 AGTTATTATTGAACGTTTAAAACAGTCTGAAATTGGTACAGGATTTAATGCGGCCAATGG 2022831

Qy   1454 TGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGC 1513

Db 2022830 AGAATGGGTAGATATGGTTACCACAGGTATCATTGACCCTGTCAAAGTAACACGTTCTGC 2022771

Qy   1514 GCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAA 1573

Db 2022770 ACTTCAAAATGCGGCATCTGTAGCAAGTCTTATCTTGACTACAGAAGCAGTAGTAGCAAA 2022711

Qy   1574 TAAACCTGAACC---AGCTACGCCAGCGCCAGCAATGCCAGCAGGTATG 1619

Db 2022710 TAAACCTGAACCAGAAGCTCCTACAGCTCCTGCAATGGATCCATCTATG 2022662


RESULT 7
ABN69159
ID   ABN69159 standard; DNA; 1620 BP.
XX
AC   ABN69159;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Streptococcus polynucleotide SEQ ID NO 6231.
XX
KW   Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW   group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW   antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS   Streptococcus agalactiae.
XX
PN   WO200234771-A2.
XX
PD   02-MAY-2002.
XX
PF   29-OCT-2001; 2001WO-GB04789.
XX
PR   27-OCT-2000; 2000GB-0026333.
PR   24-NOV-2000; 2000GB-0028727.
PR   07-MAR-2001; 2001GB-0005640.
XX
PA   (CHIR-) CHIRON SPA.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Telford J,  Masignani V,  Margarit Ros YI,  Grandi G,  Fraser C;
PI   Tettelin H;
XX
DR   WPI; 2002-352536/38.
DR   P-PSDB; ABP28528.
XX
PT   New Streptococcus protein for the treatment or prevention of infection
PT   or disease caused by Streptococcus bacteria, such as meningitis, and
PT   for detecting a compound that binds to the protein -
XX
PS   Claim 7; Page 3785; 4525pp; English.
XX
CC   The invention relates to a protein (ABP25413-ABP30895) from group B
CC   streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC   (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
CC   the specification. The proteins have antibacterial and antiinflammatory
CC   activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC   antibodies that bind (I) are used in the manufacture of medicaments for
CC   the treatment or prevention of infection or disease caused by
CC   Streptococcus bacteria, particularly S. agalactiae and S. pyrogenes.
CC   Nucleic acids encoding (I) are used to detect Streptococcus in a
CC   biological sample. (I) is used to determine whether a compound binds to
CC   (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC   used as a vaccine or diagnostic composition. The disease caused by
CC   Streptococcus that is prevented or treated may be meningitis. Nucleic
CC   acid encoding (I) may be used to recombinantly produce (I) and may be
CC   used in gene therapy. Antibodies to (I) are used for affinity
CC   chromatography, immunoassays, and distinguishing/identifying
CC   Streptococcus proteins.
XX
SQ   Sequence 1620 BP; 498 A; 271 C; 390 G; 461 T; 0 other;

  Query Match             64.0%;  Score 1062.4;  DB 24;  Length 1620;
```

```
  Best Local Similarity   79.2%;  Pred. No. 2.4e-268;
  Matches 1274;  Conservative    0;  Mismatches  331;  Indels    3;  Gaps    1;

Qy      15 ATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTGAT  74
           |||||||||||| || ||||||||||||||||||  |  || ||||||||  || ||||||
Db       1 ATGGCAAAAGATATTAAATTTTCAGCAGATGCCCGCTCAGCAATGGTGCGTGGTGTTGAT  60

Qy      75 ATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTGAA  134
           || ||||| ||||| ||||||||||||  | ||||||||||| || ||||||||||||||||
Db      61 ATTTTAGCTGATACAGTCAAAGTAACATTAGGTCCTAAAGGCCGTAATGTTGTTCTTGAA  120

Qy     135 AAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCGAA  194
           ||||| |||||||||||| |||||||| ||||| || || || ||||||||||| || ||
Db     121 AAAGCATTTGGTTCTCCTTTAATTACAAATGATGGTGTGACAATTGCTAAAGAAATTGAG  180

Qy     195 TTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAACC  254
           |||||||||| |||||||| |||||||| ||| | ||||| ||||||||||| ||||||
Db     181 CTAGAAGATCACTTTGAAAATATGGGAGCTAAACTTGTGTCAGAAGTGGCTTCAAAAACT  240

Qy     255 AATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCAT  314
           |||||||||| || ||||| || || ||||| || |||||||| ||||| ||||| |
Db     241 AATGATATTGCAGGGGATGGCACTACAACTGCTACTGTTTTGACCCAAGCTATTGTACGG  300

Qy     315 GAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGAA  374
           ||||| || |||||||| || ||||| || ||||| ||||| || ||||| || ||||||
Db     301 GAAGGTCTTAAAAATGTAACTGCAGGGGCAAATCCGATTGGCATTCGTCGTGGTATTGAA  360

Qy     375 ACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAAG  434
           |||||      || |||| |||||||   | ||||  ||||| ||||| || || |||||
Db     361 ACAGCTGTTTCAGCAGCAGTTGAAGAGCTAAAAGAGATTGCACAACCAGTTTCAGGCAAA  420

Qy     435 GAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATATC  494
           |||||||||||||| || || || || || ||||| ||||||||||||| || || ||
Db     421 GAAGCTATTGCTCAAGTTGCGGCTGTGTCTTCACGTTCTGAAAAAGTTGGGGAATACATT  480

Qy     495 TCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGT  554
           || |||||||||||||| ||||| || |||||||| || || || |||||||| ||||||
Db     481 TCTGAAGCTATGGAGCGCGTGGGTAATGATGGTGTTATCACTATTGAAGAATCGCGAGGT  540

Qy     555 ATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAA  614
           |||||||||| |||||||| || ||||| ||||| |||||||||| ||| |||| ||
Db     541 ATGGAAACAGAGCTTGAAGTTGTGGAAGGAATGCAGTTTGACCGTGGGTACTTGTCACAG  600

Qy     615 TACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATC  674
           || ||||| || || || || ||||||||  | || ||||| || || | |||| | ||
Db     601 TATATGGTAACTGATAACGAGAAAATGGTCTCTGAACTTGAGAATCCGTATATCCTTATT  660

Qy     675 ACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAA  734
           || ||||| ||| | ||||| |||||||| ||||| ||| || | || || |||||||||
Db     661 ACAGATAAGAAAATTTCAAATATCCAAGAAATTTTACCATTATTAGAAGAGGTTCTTAAA  720

Qy     735 ACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCTT  794
           || || ||||| || || || || || ||||||||| ||||| ||||| || ||||| |||
Db     721 ACAAATCGTCCGTTGCTAATCATCGCTGATGATGTTGATGGAGAAGCTCTCCCAACGCTT  780

Qy     795 GTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGT  854
           ||  | ||||| |||||||| |||||||||| || ||||| |||||||| |||||||||
Db     781 GTTCTTAACAAAATTCGTGGAACTTTCAATGTCGTAGCTGTTAAAGCGCCTGGATTTGGT  840

Qy     855 GATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACA  914
           |||||||||||||| ||||| ||||| ||||||||| | ||||| || || ||  ||||
Db     841 GATCGTCGTAAAGCCATGCTGGAAGATATTGCTATCCTAACAGGAGGAACTGTCGTTACT  900

Qy     915 GAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAG  974
           || || || ||  | || ||||||||||||||| |||  ||   | |||||| |||||||
Db     901 GAAGACCTTGGTTTAGACTTAAAAGATGCTACTATGCAAGTTTTAGGACAGTCTGCTAAA  960

Qy     975 ATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATT 1034
           | ||||| |||||||||   || || ||||| |||||| | ||    |||  ||| |||
Db     961 GTAACAGTAGATAAAGATTCTACTGTTATTGTCGAAGGTGCCGGTGACTCATCAGCAATT 1020

Qy    1035 GCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGT 1094
           ||||| ||  | ||  | ||||| || ||  | ||  | |||||||||| ||||| |||

Db    1021 GCTAATCGCGTAGCTATCATTAAGTCACAGATGGAGGCTACAACTTCTGATTTTGATCGT 1080

Qy    1095 GAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGA 1154
           |||||| ||||||||||  | || || ||||| |||||||||| || || ||||| ||
Db    1081 GAAAAATTACAAGAACGACTTGCTAAGTTAGCCGGTGGTGTAGCAGTAATTAAAGTTGGT 1140

Qy    1155 GCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCT 1214
           ||  | || || |||| |||||||| |||||||||||||||| || |||||  |||||||
Db    1141 GCAGCGACTGAAACAGAATTAAAAGAGATGAAACTTCGCATCGAAGATGCGTTAAATGCA 1200

Qy    1215 ACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTT 1274
           || |||| || |||||||||||||| ||| | |||| || || || ||| | |  |||
Db    1201 ACGCGTGCTGCAGTTGAAGAAGGTATTGTTTCAGGTGGAGGTACGGCTCTTGTGAACGTT 1260

Qy    1275 ATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGTG 1334
           |||||||||||||| || ||  | ||| | || |||| |  ||||||||||| |||||
Db    1261 ATTGAAAAAGTAGCGGCACTGAAACTTAATGGTGATGAGGAGACTGGACGTAATATTGTT 1320

Qy    1335 CTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCC 1394
           |||||||||| |||||||||| ||||||||||||  |||||||| || ||||| ||
Db    1321 CTTCGTGCTCTCGAAGAGCCTGTTCGTCAAATTGCTTACAATGCTGGATATGAAGGTTCA 1380

Qy    1395 GTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGT 1454
           ||  |||||||     || ||| |          || ||||||||||||| || |  ||
Db    1381 GTTATTATTGAACGTTTAAAACAGTCTGAAATTGGTACAGGATTTAATGCGGCCAATGGA 1440

Qy    1455 GAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCG 1514
           || ||||| |||||| |||  ||||| |||||||||||||||||||||||||||| || ||
Db    1441 GAATGGGTAGATATGGTTACCACAGGTATCATTGACCCTGTCAAAGTAACACGTTCTGCA 1500

Qy    1515 CTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAAT 1574
           |||||||||| || |||||||| |||||||| |||| |||||||||| || || |||
Db    1501 CTTCAAAATGCGGCATCTGTAGCAAGTCTTATCTTGACTACAGAAGCAGTAGTAGCAAAT 1560

Qy    1575 AAACCTGAACC---AGCTACGCCAGCGCCAGCAATGCCAGCAGGTATG 1619
           |||||||||||   |||| |  |||| || ||||||    ||  ||||
Db    1561 AAACCTGAACCAGAAGCTCCTACAGCTCCTGCAATGGATCCATCTATG 1608


RESULT 8
AAX86153
ID   AAX86153 standard; DNA; 1654 BP.
XX
AC   AAX86153;
XX
DT   22-SEP-1999  (first entry)
XX
DE   DNA encoding a Streptococcus pneumoniae heat shock protein (Hsp)60-2.
XX
KW   Heat shock protein; Hsp60-2; immune response; immunological carrier;
KW   cancer control; tumour; sarcoma; cancer; gene therapy; ss.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO9935270-A1.
XX
PD   15-JUL-1999.
XX
PF   29-DEC-1998;   98WO-CA01203.
XX
PR   31-DEC-1997;   97US-0001737.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Mizzen L,  Wisniewski J;
XX
DR   WPI; 1999-430397/36.
DR   P-PSDB; AAY23902.
XX
PT   New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT   useful in vaccines, as carriers for other immunogens, as anticancer
PT   agents and for diagnosis
XX
```

PS Claim 3; Fig 2A-B; 176pp; English.
XX
CC The present sequence encodes a heat shock protein, designated Hsp60-2.
CC The protein, its fragments, variants and fusion proteins, are
CC used to elicit or enhance an immune response against Streptococcus,
CC and to elicit a similar response to a target antigen fused to the
CC protein. Unlike other immunological carriers, Hsp60 proteins are not
CC immunosuppressive so provide an increased response to any conjugated or
CC fused antigen. Also, where used for cancer control, they lack the side
CC effects associated with endotoxins. They can also be used to detect
CC specific antibodies and in treatment or prevention of tumours
CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or
CC liver). The Hsp60 polynucleotide is used for recombinant production
CC of the protein, as a source of primers and probes for detecting
CC streptococci in standard hybridization/amplification assays, and
CC therapeutically in gene therapy vectors.
XX
SQ Sequence 1654 BP; 500 A; 326 C; 401 G; 427 T; 0 other;

Query Match 63.3%; Score 1052; DB 20; Length 1654;
Best Local Similarity 77.7%; Pred. No. 1.3e-265;
Matches 1271; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

```
Qy      1 GAATTCGGCTTCATATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGG 60
          ||||||||||||||||||||||||||||||| |||||||| ||||||| ||| | || ||||
Db      1 GAATTCGGCTTCATATGGCAAAAGAAATTAAATTTTCATCAGATGCCCGTTCAGCTATGG 60

Qy     61 TGCGCGGAGTTGATATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCA 120
          | || || || |||||  | |||||||| || |||||||| | || || ||||| ||||
Db     61 TCCGTGGTGTCGATATCCTTGCAGATACTGTTAAAGTAACTTTGGGACCAAAAGGTCGCA 120

Qy    121 ATGTTGTTCTTGAAAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTG 180
          |||| |||||||||||  | || ||||| ||||| ||||| |||||||| || || ||||
Db    121 ATGTCGTTCTTGAAAAGTCATTCGGTTCACCCTTGATTACCAATGACGGTGTGACTATTG 180

Qy    181 CTAAAGAGATCGAATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAG 240
          | ||||| || |||||||||| |||||||||| ||||| || |||||||| || ||||
Db    181 CCAAAGAAATTGAATTAGAAGACCATTTTGAAAATATGGGTGCCAAATTGGTATCAGAAG 240

Qy    241 TGGCTTCTAAAACCAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACAC 300
          | ||||| |||||||||||||| || |||||||| || || |||||||| |||||||| |
Db    241 TAGCTTCAAAAACCAATGATATCGCAGGTGATGGAACTACAACTGCAACTGTTTTGACCC 300

Qy    301 AAGCCATTGTTCATGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCC 360
          |||| || || | ||||||| | ||||| ||.|||||||| |||||||| ||||| |
Db    301 AAGCAATCGTCCGTGAAGGAATCAAAAACGTCACAGCAGGTGCAAATCCAATCGGTATTC 360

Qy    361 GTCGAGGCATTGAAACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAAC 420
          |||| || ||||||||||||   ||  |||| |||||||| |||||| ||  |    |
Db    361 GTCGTGGGATTGAAACAGCAGTTGCCGCAGCAGTTGAAGCTTTGAAAAACAACGTCATCC 420

Qy    421 CTGTATCTGGCAAGGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAG 480
          ||||  |    || |||||||| ||||| || || || ||||| || || ||||||||||
Db    421 CTGTTGCCAATAAAGAAGCTATCGCTCAAGTTGCAGCCGTATCTTCTCGTTCTGAAAAAG 480

Qy    481 TTGGAGAGTATATCTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCG 540
          |||| ||||| ||||| ||||| |||||    || ||||| || ||||| || |||||||
Db    481 TTGGTGAGTACATCTCTGAAGCAATGGAAAAAGTTGGCAAAGACGGTGTCATCACCATCG 540

Qy    541 AAGAATCTCGAGGTATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTG 600
          |||| || || |||||||||||||| |||||||| || ||||| ||||| ||||||||||
Db    541 AAGAGTCACGTGGTATGGAAACAGAGCTTGAAGTCGTAGAAGGAATGCAGTTTGACCGTG 600

Qy    601 GTTACCTGTCTCAATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACC 660
          ||||||| || || |||||||| ||||| |  |||||||||| || ||||||||||  |
Db    601 GTTACCTTTCACAGTACATGGTGACAGATAGCGAAAAAATGGTGGCTGACCTTGAAAATC 660

Qy    661 CATTTATCTTAATCACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTG 720
          | |  || || || || || || ||| | || || |||||||| || ||||||||  | |
Db    661 CGTACATTTTGATTACAGACAAGAAAATTTCCAATATCCAAGAAATCTTGCCACTTTTGG 720

Qy    721 AGGAAGTTCTTAAAACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAG 780
          |     ||||  ||| ||| |||||| |  | |||||||| ||||||||||||||||| |
Db    721 AAAGCATTCTCCAAAGCAATCGTCCACTCTTGATTATTGCGGATGATGTGGATGGTGAGG 780

Qy    781 CACTTCCAACCCTTGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAG 840
          | |||||||| ||||| |||||||||||||||||| || ||||| || || || ||||| |
Db    781 CTCTTCCAACTCTTGTTTTGAACAAGATTCGTGGAACCTTCAACGTAGTAGCAGTCAAGG 840

Qy    841 CGCCAGGATTTGGTGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTG 900
          | || || |||||||| ||||| ||||| |||||||||| || || ||||| ||||| |
Db    841 CACCTGGTTTTGGTGACCGTCGCAAAGCCATGCTTGAAGATATCGCCATCTTAACAGGCG 900

Qy    901 GTACAGTGATTACAGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTG 960
          | ||||| || ||||| || || || ||||| || |||||||| |||||   ||| ||||
Db    901 GAACAGTTATCACAGAAGACCTTGGTCTTGAGTTGAAAGATGCGACAATTGAAGCTCTTG 960

Qy    961 GACAGGCTGCTAAGATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAA 1020
          | || || || |   | || || || ||||||||||| || ||||| |||||| ||||||
Db    961 GTCAAGCAGCGAGAGTGACCGTGGACAAAGATAGCACGGTTATTGTAGAAGGTGCAGGAA 1020

Qy   1021 GTTCAGAAGCTATTGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTT 1080
          | | ||||| ||| || ||||| ||||  | || || || ||| | ||||| |||||||
Db   1021 ATCCTGAAGCGATTTCTCACCGTGTTGCGGTTATCAAGTCTCAAATCGAAACTACAACTT 1080

Qy   1081 CTGACTTTGACCGTGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTG 1140
          |||| |||||||||||||||| | |||||||| ||||| |||||  | ||||||||||| |
Db   1081 CTGAATTTGACCGTGAAAAATTGCAAGAACGCTTGGCCAAATTGTCAGGTGGTGTAGCGG 1140

Qy   1141 TTATCAAAGTAGGAGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGG 1200
          |||| || || ||||| |||| || || | || |||||||||||||| ||||||| |
Db   1141 TTATTAAGGTCGGAGCCGCAACTGAAACTGAGTTGAAAGAAATGAAACTCCGCATTGAAG 1200

Qy   1201 ATGCTCTAAATGCTACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAG 1260
          |||| || || ||||| |||||||| |||||||||||||| ||||| ||||||||||||
Db   1201 ATGCCCTCAACGCTACTCGTGCAGCTGTTGAAGAAGGTATTGTTGCAGGTGGTGGAACAG 1260

Qy   1261 CACTTATTACGGTTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTG 1320
          | |||   | || ||| |  || || |  | || |    || ||||| || || |
Db   1261 CTCTTGCCAATGTGATTCCAGCTGTTGCTACCTTGGAATTGACAGGAGATGAAGCAACAG 1320

Qy   1321 GACGTAACATTGTGCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTG 1380
          ||||||| ||||| || |||||| | ||||| ||||| |||||||||||    ||||| |
Db   1321 GACGTAATATTGTTCTCCGTGCTTTGGAAGAACCTGTTCGTCAAATTGCTCACAATGCAG 1380

Qy   1381 GGTACGAAGGCTCCGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTA 1440
          | |  ||||| ||  | ||||| ||    |||||||||           || | |||||| |
Db   1381 GATTTGAAGGATCTATCGTTATCGATCGTTTGAAAAATGCTGAGCTTGGTATAGGATTCA 1440

Qy   1441 ATGCTGCAACAGGTGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAG 1500
          | || ||||| || |||||||| | |||||| |   ||| |||||||| || || ||||
Db   1441 ACGCAGCAACTGGCGAGTGGGTTAACATGATTGATCAAGGTATCATTGATCCAGTTAAAG 1500

Qy   1501 TAACACGATCAGCGCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAG 1560
          | |  || ||||| || |||||||||||| |||||||| ||  | ||||||||||||||||
Db   1501 TGAGTCGTTCAGCCCTACAAAATGCAGCATCTGTAGCCAGCTTGATTTTGACAACAGAAG 1560

Qy   1561 CAGTTGTTGCTAATAAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGG 1620
          |||| || || |||||||| |||||||   | ||||| |||||||||    || |||||
Db   1561 CAGTCGTAGCCAATAAACCAGAACCAGTAGCCCCAGCTCCAGCAATGGATCCAAGTATGA 1620

Qy   1621 ATCCAGGAATGATGGG 1636
               ||||||   ||
Db   1621 TGGGTGGAATGGGCGG 1636
```

RESULT 9
AAH90800
ID AAH90800 standard; DNA; 1623 BP.
XX
AC AAH90800;
XX
DT 02-OCT-2001 (first entry)
XX
DE CFE 104 coding sequence.
XX

KW Antibacterial; vaccine; gene therapy; bacterial cell wall viability;
KW CFE; CEG; Conserved Essential Gene; bacterial infection;
KW antisense therapy; antibiotic resistance; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200149721-A2.
XX
PD 12-JUL-2001.
XX
PF 29-DEC-2000; 2000WO-US35604.
XX
PR 30-DEC-1999; 99US-0174089.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Dougherty TJ, Pucci MJ, Dougherty BA, Davison DB, Bruccoleri RE;
PI Thanassi JA;
XX
DR WPI; 2001-496721/54.
DR P-PSDB; AAM01101.
XX
PT Nucleic acids encoding conserved essential genes involved in bacterial
PT replication which are potential targets for the treatment of antibiotic
PT resistant bacterial infections -
XX
PS Claim 16; Page 231; 380pp; English.
XX
CC The present invention relates to nucleic acids (AAH90701-AAH90918)
CC encoding polypeptides (AAM01002-AAM01114), which are essential for the
CC viability of a bacterial cell wall. The acronym CFE stands for "CEG For
CC Expression", where CEG stands for "Conserved Essential Gene". The nucleic
CC acids are useful for detecting the presence of proteins essential for the
CC viability of a bacterial cell wall in samples such as cells, tissues,
CC biological fluids, blood, serum, nose, ear or throat swabs with ligands,
CC and for detecting corresponding target nucleic acid molecules with
CC complementary sequences. The nucleic acids are also useful for
CC determining whether a genomic nucleotide sequence of interest is
CC essential for viability of a bacterial cell or whether it resides within
CC an operon, by integrating an exogeneous nucleotide sequence comprising a
CC portion of an open reading frame of the genomic sequence of interest
CC (comprising 200-500 base pairs) into the genomic sequence of interest
CC which confers a selectable phenotype to the cell, and determining cell
CC viability with a selection agent such as chloramphenicol. The nucleic
CC acids and proteins are also useful as vaccines and for treating bacterial
CC infections with gene therapy and antisense therapy. The nucleic acids
CC also enable identification of targets suitable for the treatment of
CC antibiotic resistant bacterial infections.
XX
SQ Sequence 1623 BP; 490 A; 325 C; 394 G; 414 T; 0 other;

Query Match 62.6%; Score 1039.2; DB 22; Length 1623;
Best Local Similarity 77.6%; Pred. No. 2.9e-262;
Matches 1257; Conservative 0; Mismatches 363; Indels 0; Gaps 0;

```
Qy      15 ATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTGAT 74
           ||| |||||||||| |||||||| ||||||| ||| | |||||||| || || || |||
Db       1 ATGTCAAAAGAAATTAAATTTTCATCAGATGCCCGTTCAGCCATGGTTCGTGGTGTCGAT 60

Qy      75 ATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTGAA 134
           ||  | ||||| || || |||||||  | || || ||||| |||||||| |||||||||
Db      61 ATTCTTGCAGACACTGTTAAAGTAACCTTGGGACCAAAAGGTCGCAATGTCGTTCTTGAA 120

Qy     135 AAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCGAA 194
           ||  | || ||||| ||||| ||||| |||||||| || |||||||| ||||| ||||||
Db     121 AAGTCATTCGGTTCACCCTTGATTACCAATGACGGTGTGACCATTGCCAAAGAAATCGAA 180

Qy     195 TTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAACC 254
           || ||||| |||||||||||| ||||| || || || || || ||||| ||||||||||||
Db     181 TTGGAAGACCATTTTGAAAATATGGGTGCTAAGTTAGTATCAGAAGTAGCTTCTAAAACC 240

Qy     255 AATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCAT 314
           |||||||| || |||||| || || || |||||||||||| ||||| |||||| || || | |
```

```
Db     241 AATGATATCGCAGGTGACGGAACTACGACTGCAACAGTCTTGACCCAAGCTATCGTCCGT 300

Qy     315 GAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGAA 374
           |||||| | ||||| || |||||||||||| |||||||| ||||| ||||| || ||||||
Db     301 GAAGGAATCAAAAACGTCACAGCAGGTGCAAATCCAATCGGTATTCGTCGTGGGATTGAA 360

Qy     375 ACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAAG 434
           ||||||   ||  |||| |||||||| ||||||  ||  ||    |||||  |   ||
Db     361 ACAGCAGTTGCCGCAGCAGTTGAAGCTTTGAAAAACAACGCCATCCCTGTTGCCAATAAA 420

Qy     435 GAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATATC 494
           |||||||| ||||| || || |||||||| || || ||||||||||||| ||||| |||
Db     421 GAAGCTATCGCTCAAGTTGCAGCAGTATCTTCTCGTTCTGAAAAAGTTGGTGAGTACATC 480

Qy     495 TCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGT 554
           || ||||| |||||    || ||||| || ||||| || |||||||||||| || ||| |||
Db     481 TCTGAAGCAATGGAAAAAGTTGGCAAAGACGGTGTCATCACCATCGAAGAGTCACGTGGT 540

Qy     555 ATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAA 614
           |||||||||||| |||||||| || ||||| ||||| |||||||||||||||||| || ||
Db     541 ATGGAAACAGAGCTTGAAGTCGTAGAAGGAATGCAGTTTGACCGTGGTTACCTTTCACAG 600

Qy     615 TACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATC 674
           |||||||| ||||| |  |||||||||||| || |||||||||| || |  || || ||
Db     601 TACATGGTGACAGATAGCGAAAAAATGGTGGCTGACCTTGAAAATCCGTACATTTTGATT 660

Qy     675 ACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAA 734
           || || || ||| | || |||||||||| || ||||||||  | ||     |||| ||
Db     661 ACAGACAAGAAAATTTCCAACATCCAAGAAATCTTGCCACTTTTGGAAAGCATTCTCCAA 720

Qy     735 ACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCTT 794
           | ||| |||||| |  | |||||||| |||||||||||||||||| || |||||||| |||
Db     721 AGCAATCGTCCACTCTTGATTATTGCGGATGATGTGGATGGTGAGGCTCTTCCAACTCTT 780

Qy     795 GTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGT 854
           || |||||||||||||||| || ||||| || || || ||||| || || || ||||||
Db     781 GTTTTGAACAAGATTCGTGGAACCTTCAACGTAGTAGCAGTCAAGGCACCTGGTTTTGGT 840

Qy     855 GATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACA 914
           || ||||| ||||| |||||||||||| || || ||||| ||||| || ||||| || |||
Db     841 GACCGTCGCAAAGCCATGCTTGAAGATATCGCCATCTTAACAGGCGGAACAGTTATCACA 900

Qy     915 GAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAG 974
           || || || || ||||| || |||||||| |||||   ||| ||||| || || || |
Db     901 GAAGACCTTGGTCTTGAGTTGAAAGATGCGACAATTGAAGCTCTTGGTCAAGCAGCGAGA 960

Qy     975 ATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATT 1034
           | || || || |||||||||| || ||||| |||||| | | ||||| |||
Db     961 GTGACCGTGGACAAAGATAGCACGGTTATTGTAGAAGGTGCAGGAAATCCTGAAGCGATT 1020

Qy    1035 GCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGT 1094
           || ||||| |||| | || || || ||| | ||||| |||||||||| |||||||||
Db    1021 TCTCACCGTGTTGCGGTTATCAAGTCTCAAATCGAAACTACAACTTCTGAATTTGACCGT 1080

Qy    1095 GAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGA 1154
           |||||| | |||||||| ||||| |||||  | |||||||||| ||||| || || |||
Db    1081 GAAAAATTGCAAGAACGCTTGGCCAAATTGTCAGGTGGTGTAGCGGTTATTAAGGTCGGA 1140

Qy    1155 GCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCT 1214
           ||  |||| ||.|| |  || |||||||||||||| |||||||| ||||| || || |||
Db    1141 GCCGCAACTGAAACTGAGTTGAAAGAAATGAAACTCCGCATTGAAGATGCCCTCAACGCT 1200

Qy    1215 ACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTT 1274
           || |||||||| |||||||||||||| ||||| |||||||||||||| |||   |  ||
Db    1201 ACTCGTGCAGCTGTTGAAGAAGGTATTGTTGCAGGTGGTGGAACAGCTCTTGCCAATGTG 1260

Qy    1275 ATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGTG 1334
           ||| |  || || |  | || |    || |||| || || |||||||| ||||
Db    1261 ATTCCAGCTGTTGCTACCTTGGAATTGACAGGAGATGAAGCAACAGGACGTAATATTGTT 1320

Qy    1335 CTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCC 1394
           || |||||| | ||||| || || |||||||||||   ||||| || |  ||||| ||
Db    1321 CTCCGTGCTTTGGAAGAACCCGTTCGTCAAATTGCTCACAATGCAGGATTTGAAGGATCT 1380
```

```
Qy   1395 GTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGT 1454
Db   1381 ATCGTTATCGATCGTTTGAAAAATGCTGAGCTTGGTATAGGATTCAACGCAGCAACTGGC 1440

Qy   1455 GAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCG 1514
Db.  1441 GAGTGGGTTAACATGATTGATCAAGGTATCATTGATCCAGTTAAAGTGAGTCGTTCAGCC 1500

Qy   1515 CTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAAT 1574
Db   1501 CTACAAAATGCAGCATCTGTAGCCAGCTTGATTTTGACAACAGAAGCAGTCGTAGCCAAT 1560

Qy   1575 AAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGATG 1634
Db   1561 AAACCAGAACCAGTAGCCCCAGCTCCAGCAATGGATCCAAGTATGATGGGCGGGATGATG 1620
```

RESULT 10
AAF25036
ID AAF25036 standard; DNA; 1926 BP.
XX
AC AAF25036;
XX
DT 30-APR-2001 (first entry)
XX
DE Nucleotide sequence of Hsp65-E7 fusion protein.
XX
KW Heat shock protein; Hsp; Th1 response; Th1 cell; CD4+ T lymphocyte cell;
KW lymphocyte; Hsp65; Hsp40; Hsp10; Hsp60; Hsp71; microbial pathogen;
KW E7 protein; ss.
XX
OS Synthetic.
OS Streptococcus pneumoniae.
OS Human papillomavirus.
XX
FH Key   Location/Qualifiers
FT CDS   1..1926
FT   /*tag= a
FT   /product= "Hsp65-E7 fusion protein"
XX
PN WO200104344-A2.
XX
PD 18-JAN-2001.
XX
PF 10-JUL-2000; 2000WO-US18828.
XX
PR 08-JUL-1999; 99US-0143757.
XX
PA (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI Siegel M, Chu NR, Mizzen LA;
XX
DR WPI; 2001-138361/14.
DR P-PSDB; AAB31619.
XX
PT Screening for compounds that stimulate Th1-like responses in CD4+ T
PT lymphocyte cells
XX
PS Example 15; Fig 15A-B; 88pp; English.
XX
CC The present sequence encodes a fusion protein comprising a
CC Streptococcus pneumoniae heat shock protein (Hsp) 65 fused to a HPV16 E7
CC protein. The fusion protein is used in the method of the invention. The
CC specification describes a method of determining whether a compound
CC stimulates a Th1-like response. Th1 cells are a subset of CD4+
CC T lymphocyte cells. The method comprises contacting naive lymphocytes
CC in vitro with a fusion protein comprising at least a fragment of Hsp,
CC and then detecting the TH1-like response exhibited by the cell sample.
CC The proteins which may be used in the method of the invention are Hsp65,
CC Hsp40, Hsp10, Hsp60, and Hsp71. The method may be used to identify
CC compounds that stimulate Th1-like responses in response to microbial
CC pathogens.

XX
SQ Sequence 1926 BP; 585 A; 382 C; 464 G; 495 T; 0 other;

Query Match 62.5%; Score 1038.4; DB 22; Length 1926;
Best Local Similarity 77.5%; Pred. No. 5e-262;
Matches 1258; Conservative 0; Mismatches 366; Indels 0; Gaps 0;

```
Qy     15 ATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTGAT 74
Db      1 ATGGCAAAAGAAATTAAATTTTCATCAGATGCCCGTTCAGCTATGGTCCGTGGTGTCGAT 60

Qy     75 ATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTGAA 134
Db     61 ATCCTTGCAGATACTGTTAAAGTAACTTTGGGACCAAAAGGTCGCAATGTCGTTCTTGAA 120

Qy    135 AAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCGAA 194
Db    121 AAGTCATTCGGTTCACCCTTGATTACCAATGACGGTGTGACTATTGCCAAAGAAATTGAA 180

Qy    195 TTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAACC 254
Db    181 TTAGAAGACCATTTTGAAAATATGGGTGCCAAATTGGTATCAGAAGTAGCTTCAAAAACC 240

Qy    255 AATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCAT 314
Db    241 AATGATATCGCAGGTGATGGAACTACAACTGCAACTGTTTTGACCCAAGCAATCGTCCGT 300

Qy    315 GAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGAA 374
Db    301 GAAGGAATCAAAAACGTCACAGCAGGTGCAAATCCAATCGGTATTCGTCGTGGGATTGAA 360

Qy    375 ACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAAG 434
Db    361 ACAGCAGTTGCCGCAGCAGTTGAAGCTTTGAAAAACAACGTCATCCCTGTTGCCAATAAA 420

Qy    435 GAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATATC 494
Db    421 GAAGCTATCGCTCAAGTTGCAGCCGTATCTTCTCGTTCTGAAAAAGTTGGTGAGTACATC 480

Qy    495 TCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGT 554
Db    481 TCTGAAGCAATGGAAAAAGTTGGCAAAGACGGTGTCATCACCATCGAAGAGTCACGTGGT 540

Qy    555 ATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAA 614
Db    541 ATGGAAACAGAGCTTGAAGTCGTAGAAGGAATGCAGTTTGACCGTGGTTACCTTTCACAG 600

Qy    615 TACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATC 674
Db    601 TACATGGTGACAGATAGCGAAAAAATGGTGGCTGACCTTGAAAATCCGTACATTTTGATT 660

Qy    675 ACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAA 734
Db    661 ACAGACAAGAAAATTTCCAATATCCAAGAAATCTTGCCACTTTTGGAAAGCATTCTCCAA 720

Qy    735 ACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCTT 794
Db    721 AGCAATCGTCCACTCTTGATTATTGCGGATGATGTGGATGGTGAGGCTCTTCCAACTCTT 780

Qy    795 GTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGT 854
Db    781 GTTTTGAACAAGATTCGTGGAACCTTCAACGTAGTAGCAGTCAAGGCACCTGGTTTTGGT 840

Qy    855 GATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACA 914
Db    841 GACCGTCGCAAAGCCATGCTTGAAGATATCGCCATCTTAACAGGCGGAACAGTTATCACA 900

Qy    915 GAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAG 974
Db    901 GAAGACCTTGGTCTTGAGTTGAAAGATGCGACAATTGAAGCTCTTGGTCAAGCAGCGAGA 960

Qy    975 ATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATT 1034
```

```
Db    961 GTGACCGTGGACAAAGATAGCACGGTTATTGTAGAAGGTGCAGGAAATCCTGAAGCGATT 1020

Qy   1035 GCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGT 1094

Db   1021 TCTCACCGTGTTGCGGTTATCAAGTCTCAAATCGAAACTACAACTTCTGAATTTGACCGT 1080

Qy   1095 GAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGA 1154

Db   1081 GAAAAATTGCAAGAACGCTTGGCCAAATTGTCAGGTGGTGTAGCGGTTATTAAGGTCGGA 1140

Qy   1155 GCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCT 1214

Db   1141 GCCGCAACTGAAACTGAGTTGAAAGAAATGAAACTCCGCATTGAAGATGCCCTCAACGCT 1200

Qy   1215 ACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTT 1274

Db   1201 ACTCGTGCAGCTGTTGAAGAAGGTATTGTTGCAGGTGGTGGAACAGCTCTTGCCAATGTG 1260

Qy   1275 ATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGTG 1334

Db   1261 ATTCCAGCTGTTGCTACCTTGGAATTGACAGGAGATGAAGCAACAGGACGTAATATTGTT 1320

Qy   1335 CTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCC 1394

Db   1321 CTCCGTGCTTTGGAAGAACCTGTTCGTCAAATTGCTCACAATGCAGGATTTGAAGGATCT 1380

Qy   1395 GTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGT 1454

Db   1381 ATCGTTATCGATCGTTTGAAAAATGCTGAGCTTGGTATAGGATTCAACGCAGCAACTGGC 1440

Qy   1455 GAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCG 1514

Db   1441 GAGTGGGTTAACATGATTGATCAAGGTATCATTGATCCAGTTAAAGTGAGTCGTTCAGCC 1500

Qy   1515 CTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAAT 1574

Db   1501 CTACAAAATGCAGCATCTGTAGCCAGCTTGATTTTGACAACAGAAGCAGTCGTAGCCAAT 1560

Qy   1575 AAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGATG 1634

Db   1561 AAACCAGAACCAGTAGCCCCAGCTCCAGCAATGGATCCAAGTATGATGGGTGGAATGGGC 1620

Qy   1635 GGTG 1638

Db   1621 GGAG 1624
```

RESULT 11
AAH56860
ID AAH56860 standard; DNA; 2107 BP.
XX
AC AAH56860;
XX
DT 06-SEP-2001 (first entry)
XX
DE Antibiotic resistant S. pneumoniae groE operon sequence SEQ ID NO:6.
XX
KW Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;
KW microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;
KW Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;
KW antibacterial; antiviral; antiproliferative; antisense therapy;
KW microbial infection; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200136625-A2.
XX
PD 25-MAY-2001.
XX
PF 20-NOV-2000; 2000WO-CA01347.
XX
PR 18-NOV-1999; 99US-0166249.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Dugourd D;
XX
DR WPI; 2001-355633/37.
XX
PT Novel antisense compounds targeting nucleic acid encoding groEL or
PT groES gene of microorganism, which hybridize with and inhibit
PT expression of the genes, useful to inhibit growth of microorganism
PT having the genes -
XX
PS Disclosure; Fig 6; 110pp; English.
XX
CC The present invention specifically claims AAH56368 to AAH56832 which are
CC antisense oligonucleotides to nucleotide sequences encoding groE. More
CC generally, antisense compounds (I) comprising antisense oligonucleotides
CC of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat
CC shock protein (HSP)60) (GL) and groES (HSP10) (GS) gene from a
CC microorganism, where the antisense compound is complementary to GL or
CC GS of a microorganism and specifically hybridises with and inhibits the
CC expression of GL or GS, is claimed. (I) have antibacterial, antiviral
CC and antiproliferative activities, and can be used in antisense therapy
CC and for inhibition of expression of groES or groEL. (I) are useful for
CC inhibiting expression of GL or GS in cells or tissues in vitro. (I) are
CC also useful for inhibiting the growth of a microorganism, or inhibiting
CC the expression of GL or GS gene in a microorganism (a bacterial cell or
CC a virus) having a GL or GS gene which involves administering to the
CC microorganism or to a cell infected with the microorganism, (I). (I) are
CC also useful for treating a mammalian pathological condition mediated by
CC the microorganisms which involves identifying a eukaryotic organism
CC having a pathological condition mediated by microorganisms having a GL
CC or GS gene and administering (I) such that the growth of microorganism
CC is inhibited. The antisense compounds are utilised for diagnostics,
CC therapeutics, prophylaxis and as research reagents and kits, e.g., to
CC prevent or delay microbial infections in humans. They are also useful as
CC molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854
CC represent PCR primers for groE sequences which are used in the
CC exemplification of the present invention. AAH56855 to AAH56870 represent
CC groE nucleotide sequence given in the present invention.
XX
SQ Sequence 2107 BP; 650 A; 400 C; 508 G; 549 T; 0 other;

Query Match 62.4%; Score 1037; DB 22; Length 2107;
Best Local Similarity 77.5%; Pred. No. 1.2e-261;
Matches 1256; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

```
Qy     14 TATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTGA 73

Db    484 TATGTCAAAAGAAATTAAATTTTCATCAGATGCCCGTTCAGCTATGGTCCGTGGTGTCGA 543

Qy     74 TATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTGA 133

Db    544 TATCCTTGCAGATACTGTTAAAGTAACCTTGGGACCAAAAGGTCGCAATGTCGTTCTTGA 603

Qy    134 AAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCGA 193

Db    604 AAAATCATTTGGCTCACCATTGATTACCAATGACGGTGTGACTATTGCCAAAGAAATTGA 663

Qy    194 ATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAAC 253

Db    664 ATTAGAAGACCATTTTGAAAATATGGGTGCCAAATTGGTATCAGAAGTAGCTTCTAAAAC 723

Qy    254 CAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCA 313

Db    724 CAATGATATCGCAGGTGACGGGACTACGACTGCAACAGTCTTGACCCAAGCTATCGTCCG 783

Qy    314 TGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGA 373

Db    784 TGAAGGAATCAAAAACGTCACAGCAGGTGCAAATCCAATCGGTATTCGTCGTGGGATTGA 843

Qy    374 AACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAA 433

Db    844 AACAGCAGTTGCCGCAGCAGTTGAAGCTTTGAAAAACAACGCCATCCCTGTTGCCAATAA 903
```

```
Qy    434 GGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATAT 493
Db    904 AGAAGCTATCGCTCAAGTTGCAGCCGTATCTTCTCGTTCTGAAAAAGTTGGTGAGTACAT 963

Qy    494 CTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGG 553
Db    964 CTCTGAAGCAATGGAAAAAGTTGGCAAAGACGGTGTCATCACCATCGAAGAGTCACGTGG 1023

Qy    554 TATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCA 613
Db   1024 TATGGAAACAGAGCTTGAAGTCGTAGAAGGAATGCAGTTTGACCGTGGTTACCTTTCACA 1083

Qy    614 ATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAAT 673
Db   1084 GTACATGGTGACAGATAGCGAAAAAATGGTGGCTGACCTTGAAAATCCGTACATTTTGAT 1143

Qy    674 CACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAA 733
Db   1144 TACAGACAAGAAAATTTCCAATATCCAAGAGATCTTGCCACTTTTGGAAAGCATTCTCCA 1203

Qy    734 AACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCT 793
Db   1204 AAGCAATCGTCCACTCTTGATTATTGCGGATGATGTGGATGGCGAGGCTCTTCCAACTCT 1263

Qy    794 TGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGG 853
Db   1264 TGTTTTGAACAAGATTCGTGGAACCTTCAACGTAGTAGCAGTCAAGGCACCTGGTTTTGG 1323

Qy    854 TGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTAC 913
Db   1324 TGACCGTCGCAAAGCCATGCTTGAAGATATCGCCATCTTAACAGGCGGAACAGTTATCAC 1383

Qy    914 AGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAA 973
Db   1384 AGAAGACCTTGGTCTTGAGTTGAAAGATGCGACAATTGAAGCTCTTGGTCAAGCAGCGAG 1443

Qy    974 GATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTAT 1033
Db   1444 AGTGACCGTGGACAAAGATAGCACGGTTATTGTAGAAGGTGCAGGAAATCCTGAAGCGAT 1503

Qy   1034 TGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCG 1093
Db   1504 TTCTCACCGTGTTGCGGTTATCAAGTCTCAAATCGAAACTACAACTTCTGAATTTGACCG 1563

Qy   1094 TGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGG 1153
Db   1564 TGAAAAATTGCAAGAACGCTTGGCCAAATTGTCAGGTGGTGTAGCGGTTATTAAGGTCGG 1623

Qy   1154 AGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGC 1213
Db   1624 AGCCGCAACTGAAACTGAGTTGAAAGAAATGAAACTCCGCATTGAAGATGCCCTCAACGC 1683

Qy   1214 TACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGT 1273
Db   1684 TACTCGTGCAGCTGTTGAAGAAGGTATTGTTGCAGGCGGTGGAACAGCTCTTGCCAATGT 1743

Qy   1274 TATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGT 1333
Db   1744 GATTCCAGCTGTTGCTACCTTGGAATTGACAGGAGATGAAGCAACAGGACGTAATATTGT 1803

Qy   1334 GCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTC 1393
Db   1804 TCTCCGTGCTTTGGAAGAACCCGTTCGTCAAATTGCTCACAATGCAGGATTTGAAGGATC 1863

Qy   1394 CGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGG 1453
Db   1864 TATCGTTATCGATCGTTTGAAAAATGCTGAGCTTGGTATAGGATTTAACGCAGCAACTGG 1923

Qy   1454 TGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGC 1513
Db   1924 CGAGTGGGTTAACATGATTGATCAAGGTATCATTGATCCAGTTAAAGTGAGTCGTTCAGC 1983

Qy   1514 GCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAA 1573
Db   1984 CCTACAAAATGCAGCATCTGTAGCCAGCTTGATTTTGACAACAGAAGCAGTCGTAGCCAA 2043

Qy   1574 TAAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGAT 1633
Db   2044 TAAACCAGAACCAGTAGCCCCAGCTCCAGCAATGGATCCAAGCATGATGGGCGGGATGAT 2103

Qy   1634 G 1634
Db   2104 G 2104
```

RESULT 12
AAH90906
ID AAH90906 standard; DNA; 1647 BP.
XX
AC AAH90906;
XX
DT 02-OCT-2001 (first entry)
XX
DE 2CFE 104 coding sequence.
XX
KW Antibacterial; vaccine; gene therapy; bacterial cell wall viability;
KW CFE; CEG; Conserved Essential Gene; bacterial infection;
KW antisense therapy; antibiotic resistance; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200149721-A2.
XX
PD 12-JUL-2001.
XX
PF 29-DEC-2000; 2000WO-US35604.
XX
PR 30-DEC-1999; 99US-0174089.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Dougherty TJ, Pucci MJ, Dougherty BA, Davison DB, Bruccoleri RE;
PI Thanassi JA;
XX
DR WPI; 2001-496721/54.
XX
PT Nucleic acids encoding conserved essential genes involved in bacterial
PT replication which are potential targets for the treatment of antibiotic
PT resistant bacterial infections -
XX
PS Claim 30; Fig 118; 380pp; English.
XX
CC The present invention relates to nucleic acids (AAH90701-AAH90918)
CC encoding polypeptides (AAM01002-AAM01114), which are essential for the
CC viability of a bacterial cell wall. The acronym CFE stands for "CEG For
CC Expression", where CEG stands for "Conserved Essential Gene". The nucleic
CC acids are useful for detecting the presence of proteins essential for the
CC viability of a bacterial cell wall in samples such as cells, tissues,
CC biological fluids, blood, serum, nose, ear or throat swabs with ligands,
CC and for detecting corresponding target nucleic acid molecules with
CC complementary sequences. The nucleic acids are also useful for
CC determining whether a genomic nucleotide sequence of interest is
CC essential for viability of a bacterial cell or whether it resides within
CC an operon, by integrating an exogeneous nucleotide sequence comprising a
CC portion of an open reading frame of the genomic sequence of interest
CC (comprising 200-500 base pairs) into the genomic sequence of interest
CC which confers a selectable phenotype to the cell, and determining cell
CC viability with a selection agent such as chloramphenicol. The nucleic
CC acids and proteins are also useful as vaccines and for treating bacterial
CC infections with gene therapy and antisense therapy. The nucleic acids
CC also enable identification of targets suitable for the treatment of
CC antibiotic resistant bacterial infections.
XX
SQ Sequence 1647 BP; 494 A; 340 C; 397 G; 416 T; 0 other;

```
Query Match           62.4%;  Score 1036;  DB 22;  Length 1647;
Best Local Similarity  77.5%;  Pred. No. 2e-261;
Matches 1255;  Conservative    0;  Mismatches  365;  Indels    0;  Gaps    0;

Qy      15 ATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTGAT 74
           ||| |||||||||| |||||||| ||||||| ||| | |||||||| || || || |||
Db       1 ATGTCAAAAGAAATTAAATTTTCATCAGATGCCCGTTCAGCCATGGTTCGTGGTGTCGAT 60

Qy      75 ATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTGAA 134
           || | ||||| || || |||||||| | || || ||||| |||||||| |||||||||
Db      61 ATCCTTGCAGACACTGTTAAAGTAACCTTGGGACCAAAAGGTCGCAATGTCGTTCTTGAA 120

Qy     135 AAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCGAA 194
           ||  | || ||||| ||||| ||||| |||||||| || |||||||| ||||| ||||||
Db     121 AAGTCATTCGGTTCACCCTTGATTACCAATGACGGTGTGACCATTGCCAAAGAAATCGAA 180

Qy     195 TTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAACC 254
           || ||||| |||||||||| ||||| || || || || || || ||||| ||||||||||||
Db     181 TTGGAAGACCATTTTGAAAATATGGGTGCTAAGTTAGTATCAGAAGTAGCTTCTAAAACC 240

Qy     255 AATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCAT 314
           |||||||| || ||||| ||||| || |||||||||||| ||||| ||||| || || | |
Db     241 AATGATATCGCAGGTGACGGGACTACGACTGCAACAGTCTTGACCCAAGCTATCGTCCGT 300

Qy     315 GAAGGACTAAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGAA 374
           |||||| | ||||| || |||||||||||| |||||||| ||||| ||||| || ||||||
Db     301 GAAGGAATCAAAAACGTCACAGCAGGTGCAAATCCAATCGGTATTCGTCGTGGGATTGAA 360

Qy     375 ACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAAG 434
           ||||||   ||  |||| || ||||| ||||||  ||  ||    |||||  |    ||
Db     361 ACAGCAGTTGCCGCAGCAGTCGAAGCTTTGAAAAACAACGCCATCCCTGTTGCCAATAAA 420

Qy     435 GAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATATC 494
           |||||||| ||||| || || || ||||| || || ||||||||||||||| ||||| |||
Db     421 GAAGCTATCGCTCAAGTTGCAGCCGTATCTTCTCGTTCTGAAAAAGTTGGTGAGTACATC 480

Qy     495 TCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGT 554
           || ||||| |||||    || ||||| || ||||| || |||||||||| || || |||
Db     481 TCTGAAGCAATGGAAAAAGTTGGCAAAGACGGTGTCATCACCATCGAAGAGTCACGTGGT 540

Qy     555 ATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAA 614
           |||||||||| |||||||| || ||||| ||||| |||||||||||||||||| || ||
Db     541 ATGGAAACAGAGCTTGAAGTCGTAGAAGGAATGCAGTTTGACCGTGGTTACCTTTCACAG 600

Qy     615 TACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATC 674
           |||||||| || || |  |||||||||| || |||||||||| || |  || || ||
Db     601 TACATGGTGACTGATAGCGAAAAAATGGTGGCTGACCTTGAAAATCCGTACATTTTGATT 660

Qy     675 ACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAA 734
           || || || ||| | || || |||||||| || ||||||||  | ||      ||||  ||
Db     661 ACAGACAAGAAAATTTCCAATATCCAAGAAATCTTGCCACTTTTGGAAAGCATTCTCCAA 720

Qy     735 ACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCTT 794
           | ||| |||||| |  | |||||||| |||||||||||||| || || |||||||| |||
Db     721 AGCAATCGTCCACTCTTGATTATTGCGGATGATGTGGATGGCGAGGCTCTTCCAACTCTT 780

Qy     795 GTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGT 854
           || |||||||||||||||| || ||||| || || || ||||| || || || |||||||
Db     781 GTTTTGAACAAGATTCGTGGAACCTTCAACGTAGTAGCAGTCAAGGCACCTGGTTTTGGT 840

Qy     855 GATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACA 914
           || ||||| ||||| |||||||||||| || || ||||| ||||| || ||||| || |||
Db     841 GACCGTCGCAAAGCCATGCTTGAAGATATCGCCATCTTAACAGGCGGAACAGTTATCACA 900

Qy     915 GAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAG 974
           || || || || ||||| || |||||||| |||||   ||| ||||| || || || |
Db     901 GAAGACCTTGGTCTTGAGTTGAAAGATGCGACAATTGAAGCTCTTGGTCAAGCAGCGAGA 960

Qy     975 ATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATT 1034
           | || || || |||||||||| || ||||| |||||| |||||| | | ||||| |||
Db     961 GTGACCGTGGACAAAGATAGCACGGTTATTGTAGAAGGTGCAGGAAATCCTGAAGCGATT 1020

Qy    1035 GCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGT 1094
           || ||||| |||| | || || || ||| | ||||| |||||||||| |||||||||
Db    1021 TCTCACCGTGTTGCGGTTATCAAGTCTCAAATCGAAACTACAACTTCTGAATTTGACCGT 1080

Qy    1095 GAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGA 1154
           |||||| | |||||||| ||||| |||||  | |||||||||| ||||| || || |||
Db    1081 GAAAAATTGCAAGAACGCTTGGCCAAATTGTCAGGTGGTGTAGCGGTTATTAAGGTTGGA 1140

Qy    1155 GCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCT 1214
           ||  |||| || || |  || |||||||||||||| |||||||| ||||| || || |||
Db    1141 GCCGCAACTGAAACTGAGTTGAAAGAAATGAAACTCCGCATTGAAGATGCCCTCAACGCT 1200

Qy    1215 ACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGTT 1274
           || |||||||| |||||||||||||| ||||| |||||||||||||| |||   |  ||
Db    1201 ACTCGTGCAGCTGTTGAAGAAGGTATTGTTGCAGGTGGTGGAACAGCTCTTGCCAATGTG 1260

Qy    1275 ATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGTG 1334
           |||  |   || ||  |  | ||  |    || ||||| || || |||||||| |||||
Db    1261 ATTCCAGCTGTTGCTACCTTGGAATTGACAGGAGATGAAGCAACAGGACGTAATATTGTT 1320

Qy    1335 CTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTCC 1394
           || |||||| | ||||| || || |||||||||||||   ||||| || |  ||||| ||
Db    1321 CTCCGTGCTTTGGAAGAACCCGTTCGTCAAATTGCTCACAATGCAGGATTTGAAGGATCT 1380

Qy    1395 GTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGGT 1454
           | ||||| ||      ||||||||         || | |||||||||| || |||||  ||
Db    1381 ATCGTTATCGATCGTTTGAAAAATGCTGAGCTTGGTATAGGATTTAACGCAGCAACTGGC 1440

Qy    1455 GAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGCG 1514
           ||||||||| | ||||||||   ||| |||||||| || || ||||| |  || |||||
Db    1441 GAGTGGGTTAACATGATTAATCAAGGTATCATTGATCCAGTTAAAGTGAGTCGTTCAGCC 1500

Qy    1515 CTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAAT 1574
           || |||||||||| |||||||| ||  | |||||||||||||||||||| || || |||
Db    1501 CTACAAAATGCAGCATCTGTAGCCAGCTTGATTTTGACAACAGAAGCAGTCGTAGCCAAT 1560

Qy    1575 AAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGATG 1634
           ||||| |||||||   | ||||| |||||||||   || |||||      || ||||||
Db    1561 AAACCAGAACCAGTAGCCCCAGCTCCAGCAATGGATCCAAGTATGATGGGCGGGATGATG 1620


RESULT 13
AAV52210
ID   AAV52210 standard; DNA; 5365 BP.
XX
AC   AAV52210;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Streptococcus pneumoniae genome fragment SEQ ID NO:77.
XX
KW   Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW   computer readable medium; vaccine; pharmaceutical composition; ds.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO9818931-A2.
XX
PD   07-MAY-1998.
XX
PF   30-OCT-1997;   97WO-US19588.
XX
PR   31-OCT-1996;   96US-0029960.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Dougherty BA,  Fannon M;
PI   Kunsch CA,  Rosen CA;
XX
DR   WPI; 1998-272225/24.
XX
PT   Computer-readable medium with recorded Streptococcus pneumoniae
PT   polynucleotide sequences - useful in diagnostic kits and assays, and
PT   pharmaceutical compositions and vaccines for Streptococcus
```

PT pneumoniae
XX
PS Claim 1; Page 628-631; 1409pp; English.
XX
CC The present invention describes a computer readable medium which has
CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
CC recorded on it, or a representative fragment or a sequence at least 95%
CC identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
CC Streptococcus pneumoniae. The present invention also describes an
CC isolated nucleic acid molecule encoding a homologue of any of the
CC fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
CC nucleic acid molecule is produced by a process comprising: (a) screening
CC a genomic DNA library using as a probe a target sequence defined by any
CC of the sequences in SEQ ID NO:1 to 391, identifying members of the
CC library which contain sequences that hybridise to the target sequence and
CC isolating the nucleic acid molecules from the members; or (b) isolating
CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
CC molecules whose nucleotide sequence is homologous to amplification
CC primers derived from the fragment of the S. pneumoniae genome to prime
CC the amplification and isolating the amplified sequences. The computer
CC readable medium can be used in a computer-based system for identifying
CC fragments of the S. pneumoniae genome of commercial importance, or
CC expression modulating fragments of the S. pneumoniae genome. Products
CC from the present invention can be used in diagnosis kits and assays, and
CC pharmaceutical compositions and vaccines for S. pneumoniae.
XX
SQ Sequence 5365 BP; 1654 A; 960 C; 1082 G; 1669 T; 0 other;

Query Match 62.3%; Score 1035.4; DB 19; Length 5365;
Best Local Similarity 77.4%; Pred. No. 4.6e-261;
Matches 1255; Conservative 0; Mismatches 366; Indels 0; Gaps 0;

```
Qy     14 TATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTGA 73
Db    277 TATGTCAAAAGAAATTAAATTTTCATCAGATGCCCGTTCAGCCATGGTTCGTGGTGTCGA 336

Qy     74 TATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTGA 133
Db    337 TATCCTTGCAGACACTGTTAAAGTAACCTTGGGACCAAAAGGTCGCAATGTCGTTCTTGA 396

Qy    134 AAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCGA 193
Db    397 AAAGTCATTCGGTTCACCCTTGATTACCAATGACGGTGTGACCATTGCCAAAGAAATCGA 456

Qy    194 ATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAAC 253
Db    457 ATTGGAAGACCATTTTGAAAATATGGGTGCTAAGTTAGTATCAGAAGTAGCTTCTAAAAC 516

Qy    254 CAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCA 313
Db    517 CAATGATATCGCAGGTGACGGAACTACGACTGCAACAGTCTTGACCCAAGCTATCGTCCG 576

Qy    314 TGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGA 373
Db    577 TGAAGGAATCAAAAACGTCACAGCAGGTGCAAATCCAATCGGTATTCGTCGTGGGATTGA 636

Qy    374 AACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAA 433
Db    637 AACAGCAGTTGCCGCAGCAGTTGAAGCTTTGAAAAACAACGCCATCCCTGTTGCCAATAA 696

Qy    434 GGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATAT 493
Db    697 AGAAGCTATCGCTCAAGTTGCAGCCGTATCTTCTCGTTCTGAAAAAGTTGGTGAGTACAT 756

Qy    494 CTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGG 553
Db    757 CTCTGAAGCAATGGAAAAAGTTGGCAAAGACGGTGTCATCACCATCGAAGAGTCACGTGG 816

Qy    554 TATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCA 613
Db    817 TATGGAAACAGAGCTTGAAGTCGTAGAAGGAATGCAGTTTGACCGTGGTTACCTTTCACA 876

Qy    614 ATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAAT 673
Db    877 GTACATGGTGACAGATAGCGAAAAAATGGTGGCTGACCTTGAAAATCCGTACATTTTGAT 936

Qy    674 CACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAA 733
Db    937 TACAGACAAGAAAATTTCCAATATCCAAGAAATCTTGCCACTTTTGGAAAGCATTCTCCA 996

Qy    734 AACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCT 793
Db    997 AAGCAATCGTCCACTCTTGATTATTGCGGATGATGTGGATGGCGAGGCTCTTCCAACTCT 1056

Qy    794 TGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGG 853
Db   1057 TGTTTTGAACAAGATTCGTGGAACCTTCAACGTAGTAGCAGTCAAGGCACCTGGTTTTGG 1116

Qy    854 TGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTAC 913
Db   1117 TGACCGTCGCAAAGCCATGCTTGAAGATATCGCCATCTTAACAGGCGGAACAGTTATCAC 1176

Qy    914 AGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAA 973
Db   1177 AGAAGACCTTGGTCTTGAGTTGAAAGATGCGACAATTGAAGCTCTTGGTCAAGCAGCGAG 1236

Qy    974 GATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTAT 1033
Db   1237 AGTGACCGTGGACAAAGATAGCACGGTTATTGTAGAAGGTGCAGGAAATCCTGAAGCGAT 1296

Qy   1034 TGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCG 1093
Db   1297 TTCTCACCGTGTTGCGGTTATCAAGTCTCAAATCGAAACTACAACTTCTGAATTTGACCG 1356

Qy   1094 TGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGG 1153
Db   1357 TGAAAAATTGCAAGAACGCTTGGCCAAATTGTCAGGTGGTGTAGCGGTTATTAAGGTTGG 1416

Qy   1154 AGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGC 1213
Db   1417 AGCCGCAACTGAAACTGAGTTGAAAGAAATGAAACTCCGCATTGAAGATGCCCTCAACGC 1476

Qy   1214 TACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGT 1273
Db   1477 TACTCGTGCAGCTGTTGAAGAAGGTATTGTTGCAGGTGGTGGAACAGCTCTTGCCAATGT 1536

Qy   1274 TATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGT 1333
Db   1537 GATTCCAGCTGTTGCTACCTTGGAATTGACAGGAGATGAAGCAACAGGACGTAATATTGT 1596

Qy   1334 GCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTC 1393
Db   1597 TCTCCGTGCTTTGGAAGAACCCGTTCGTCAAATTGCTCACAATGCAGGATTTGAAGGATC 1656

Qy   1394 CGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGG 1453
Db   1657 TATCGTTATCGATCGTTTGAAAAATGCTGAGCTTGGTATAGGATTTAACGCAGCAACTGG 1716

Qy   1454 TGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGC 1513
Db   1717 CGAGTGGGTTAACATGATTGATCAAGGTATCATTGATCCAGTTAAAGTGAGTCGTTCAGC 1776

Qy   1514 GCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAA 1573
Db   1777 CCTACAAAATGCAGCATCTGTAGCCAGCTTGATTTTGACAACAGAAGCAGTCGTAGCCAA 1836

Qy   1574 TAAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGAT 1633
Db   1837 TAAACCAGAACCAGTAGCCCCAGCTCCAGCAATGGATCCAAGCATGATGGGCGGGATGAT 1896

Qy   1634 G 1634
Db   1897 G 1897
```

RESULT 14
AAH56859
ID AAH56859 standard; DNA; 2401 BP.
XX
AC AAH56859;
XX
DT 06-SEP-2001 (first entry)
XX
DE S. pneumoniae groE operon nucleotide sequence SEQ ID NO:5.
XX
KW Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;
KW microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;
KW Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;
KW antibacterial; antiviral; antiproliferative; antisense therapy;
KW microbial infection; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200136625-A2.
XX
PD 25-MAY-2001.
XX
PF 20-NOV-2000; 2000WO-CA01347.
XX
PR 18-NOV-1999; 99US-0166249.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Dugourd D;
XX
DR WPI; 2001-355633/37.
XX
PT Novel antisense compounds targeting nucleic acid encoding groEL or
PT groES gene of microorganism, which hybridize with and inhibit
PT expression of the genes, useful to inhibit growth of microorganism
PT having the genes -
XX
PS Disclosure; Fig 5; 110pp; English.
XX
CC The present invention specifically claims AAH56368 to AAH56832 which are
CC antisense oligonucleotides to nucleotide sequences encoding groE. More
CC generally, antisense compounds (I) comprising antisense oligonucleotides
CC of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat
CC shock protein (HSP)60) (GL) and groES (HSP10) (GS) gene from a
CC microorganism, where the antisense compound is complementary to GL or
CC GS of a microorganism and specifically hybridises with and inhibits the
CC expression of GL or GS, is claimed. (I) have antibacterial, antiviral
CC and antiproliferative activities, and can be used in antisense therapy
CC and for inhibition of expression of groES or groEL. (I) are useful for
CC inhibiting expression of GL or GS in cells or tissues in vitro. (I) are
CC also useful for inhibiting the growth of a microorganism, or inhibiting
CC the expression of GL or GS gene in a microorganism (a bacterial cell or
CC a virus) having a GL or GS gene which involves administering to the
CC microorganism or to a cell infected with the microorganism, (I). (I) are
CC also useful for treating a mammalian pathological condition mediated by
CC the microorganisms which involves identifying a eukaryotic organism
CC having a pathological condition mediated by microorganisms having a GL
CC or GS gene and administering (I) such that the growth of microorganism
CC is inhibited. The antisense compounds are utilised for diagnostics,
CC therapeutics, prophylaxis and as research reagents and kits, e.g., to
CC prevent or delay microbial infections in humans. They are also useful as
CC molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854
CC represent PCR primers for groE sequences which are used in the
CC exemplification of the present invention. AAH56855 to AAH56870 represent
CC groE nucleotide sequence given in the present invention.
XX
SQ Sequence 2401 BP; 743 A; 459 C; 568 G; 631 T; 0 other;

Query Match 62.0%; Score 1030.6; DB 22; Length 2401;
Best Local Similarity 77.2%; Pred. No. 6.1e-260;
Matches 1252; Conservative 0; Mismatches 369; Indels 0; Gaps 0;

```
Qy     14 TATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGTTGA 73
Db    618 TATGTCAAAAGAAATTAAATTTTCATCAGATGCCCGTTCAGCCATGGTTCGTGGTGTCGA 677
Qy     74 TATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTGA 133
Db    678 TATCCTTGCAGACACTGTTAAAGTAACCTTGGGACCAAAAGATCGCAATGTCGTTCTTGA 737
Qy    134 AAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCGA 193
Db    738 AAAGTCATTCGGTTCACCCTTGATTACCAATGACGGTGTGACCATTGCCAAAGAAATCGA 797
Qy    194 ATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAAAAC 253
Db    798 ATTGGAAGACCATTTTGAAAATATGGGTGCTAAGTTAGTATCAGAAATAGCTTCTAAAAC 857
Qy    254 CAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGTTCA 313
Db    858 CAATGATATCGCAGGTGACGGGACTACGACTGCAACAGTCTTGACCCAAGCTATCGTCCG 917
Qy    314 TGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCATTGA 373
Db    918 TGAAGGAATCAAAAACGTCACAGCAGGTGCAAATCCAATCGGTATTCGTCGTGGGATTGA 977
Qy    374 AACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAA 433
Db    978 AACAGCAGTTGCCGCAGCAGTTGAAGCTTTGAAAAACAACGCCATCCCTGTTGCCAATAA 1037
Qy    434 GGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATAT 493
Db   1038 AGAAGCTATCTCTCAAGTTGCAGCCGTATCTTCTCGTTCTGAAAAAGTTGGTGAGTACAT 1097
Qy    494 CTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGG 553
Db   1098 CTCTGAAGCAATGGAAAAAGTTGGCAAAGACGGTGTCATCACCATCGAAGAGTCACGTGG 1157
Qy    554 TATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCA 613
Db   1158 TATGGAAACAGAGCTTGAAGTCGTAGAAGGAATGCAGTTTGACCGTGGTTACCTTTCACA 1217
Qy    614 ATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAAT 673
Db   1218 GTACATGGTGACTGATAGCGAAAAAATGGTGGCTGACCTTGAAAATCCGTACATTTTGAT 1277
Qy    674 CACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAA 733
Db   1278 TACAGACAAGAAAATTTCCAATATCCAAGAAATCTTGCCACTTTTGGAAAGCATTCTCCA 1337
Qy    734 AACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAACCCT 793
Db   1338 AAGCAATCGTCCACTCTTGATTATTGCGGATGATGTGGATGGCGAGGCTCTTCCAACTCT 1397
Qy    794 TGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGG 853
Db   1398 TGTTTTGAACAAGATTCGTGGAACCTTCAACGTAGTAGCAGTCAAGGCACCTGGTTTTGG 1457
Qy    854 TGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTAC 913
Db   1458 TGACCGTCGCAAAGCCATGCTTGAAGATATCGCCATCTTAACAGGCGGAACAGTTATCAC 1517
Qy    914 AGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAA 973
Db   1518 AGAAGACCTTGGTCTTGAGTTGAAAGATGCGACAATTGAAGCTCTTGGTCAAGCAGCGAG 1577
Qy    974 GATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTAT 1033
Db   1578 AGTGACCGTGGACAAAGATAGCACGGTTATTGTAGAAGGTGCAGGAAATCCTGAAGCGAT 1637
Qy   1034 TGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCG 1093
Db   1638 TTCTCACCGTGTTGCGGTTATCAAGTCTCAAATCGAAACTACAACTTCTGAATTTGACCG 1697
Qy   1094 TGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGG 1153
```

```
Db   1698 TGAAAAATTGCAAGAACGCTTGGCCAAATTGTCAGGTGGTGTAGCGGTTATTAAGGTTGG 1757

Qy   1154 AGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGC 1213

Db   1758 AGCCGCAACTGAAACTGAGTTGAAAGAAATGAAACTCCGCATTGAAGATGCCCTCAACGC 1817

Qy   1214 TACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTACGGT 1273

Db   1818 TACTCGTGCAGCTGTTGAAGAAGGTATTGTTGCAGGTGGTGGAACAGCTCTTGCCAATGT 1877

Qy   1274 TATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACATTGT 1333

Db   1878 GATTCCAGCTGAAGCTACCTTGGAATTGACAGGAGATGAAGCAACAGGACGTAATATTGT 1937

Qy   1334 GCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGGCTC 1393

Db   1938 TCTCCGTGCTTTGGAAGAACCCGTTCGTCAAATTGCTCACAATGCAGGATTTGAAGGATC 1997

Qy   1394 CGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACAGG 1453

Db   1998 TATCGTTATCGATCGTTTGAAAAATGCTGAGCTTGGTATAGGATTTAACGCAGCAACTGG 2057

Qy   1454 TGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATCAGC 1513

Db   2058 CGAGTGGGTTAACATGATTGATCAAGGTATCATTGATCCAGTTAAAGTGAGTCGTTCAGC 2117

Qy   1514 GCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGCTAA 1573

Db   2118 CCTACAAAATGCAGCATCTGTAGCCAGCTTGATTTTGACAACAGAAGCAGTCGTAGCCAA 2177

Qy   1574 TAAACCTGAACCAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGGATCCAGGAATGAT 1633

Db   2178 TAAACCAGAACCAGTAGCCCCAGCTCCAGCAATGGATCCAAGCATGATGGGCGGGATGAT 2237

Qy   1634 G 1634

Db   2238 G 2238


RESULT 15
ABA90521
ID   ABA90521 standard; DNA; 2365589 BP.
XX
AC   ABA90521;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Genomic sequence of Lactococcus lactis IL1403.
XX
KW   Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.
XX
OS   Lactococcus lactis IL1403.
XX
PN   FR2807446-A1.
XX
PD   12-OCT-2001.
XX
PF   11-APR-2000; 2000FR-0004630.
XX
PR   11-APR-2000; 2000FR-0004630.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI   Bolotine A,  Sorokine A,  Renault P,  Ehrlich SD;
XX
DR   WPI; 2002-043418/06.
XX
PT   New nucleotide sequence useful in the identification or Lactococcus
PT   lactis and related species -
XX
PS   Claim 1; SEQ ID 1; 2504pp; French.
XX
CC   The present invention is related to a Lactococcus lactis nucleotide
CC   sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC   nucleic acid sequence is useful in the detection and/or amplification of
CC   nucleic acid sequence, particularly to identify Lactococcus lactis or
CC   related species. The proteins of the invention are useful for the
CC   biosynthesis or biodegradation of a composition of interest. The
CC   invention helps research in lactic bacteria, particularly useful in the
CC   production of yogurt and cheese.
CC   Note: The sequence data for this patent is based on equivalent patent
CC   WO200177334 (published 18-OCT-2001) which is available in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2365589 BP; 765914 A; 415261 C; 420487 G; 763927 T; 0 other;

  Query Match             57.5%;  Score 955;  DB 24;  Length 2365589;
  Best Local Similarity   75.1%;  Pred. No. 6.4e-239;
  Matches 1192;  Conservative    0;  Mismatches  395;  Indels    0;  Gaps    0;

Qy      11 TCATATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCTGCCATGGTGCGCGGAGT 70

Db 398893 TAAAATGTCAAAAGATATTAAATTTTCAAGTGATGCTAGAACAGCGATGATGCGTGGGAT 398952

Qy      71 TGATATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCT 130

Db 398953 TGATATTCTTGCTGATACAGTAAAAACAACCCTTGGACCAAAAGGTCGCAATGTTGTTCT 399012

Qy     131 TGAAAAAGCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGAT 190

Db 399013 TGAAAAATCATATGGATCACCTTTAATTACAAACGATGGTGTAACGATTGCCAAAGAGAT 399072

Qy     191 CGAATTAGAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTGAAGTGGCTTCTAA 250

Db 399073 TGAACTTGAAGATCATTTTGAAAATATGGGAGCTAAACTTGTTTCAGAAGTTGCTTCAAA 399132

Qy     251 AACCAATGATATTGCTGGTGATGGGACGACTACTGCAACAGTTTTGACACAAGCCATTGT 310

Db 399133 AACAAATGATATCGCAGGTGACGGTACAACAACAGCGACAGTTTTGACACAAGCTATCGT 399192

Qy     311 TCATGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGAGGCAT 370

Db 399193 TCGTGAAGGTTTAAAAAATGTTACCGCAGGTGCAAATCCAGTTGGTATTCGCCGAGGAAT 399252

Qy     371 TGAAACAGCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGG 430

Db 399253 TGAACTTGCTGCTGAAACAGCTGTTGCGTCAATTAAAGAGATGGCAATTCCTGTCCACGA 399312

Qy     431 CAAGGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAAGTTGGAGAGTA 490

Db 399313 TAAATCAGCAATTGCGCAAGTAGCTACCGTTTCATCACGTAGTGAAAAAGTGGGTGAGTA 399372

Qy     491 TATCTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCG 550

Db 399373 TATTTCTGATGCAATGGAACGTGTAGGTTCTGACGGAGTTATCACCATTGAAGAATCAAA 399432

Qy     551 AGGTATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTC 610

Db 399433 AGGGATGCAAACTGAACTCGATGTTGTTGAAGGAATGCAATTTGACCGTGGATATTTGAG 399492

Qy     611 TCAATACATGGTCACAGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTT 670

Db 399493 TCAATATATGGTTTCTAATACAGAAAAAATGGTTGCTGAATTAGATAATCCTTATATTCT 399552

Qy     671 AATCACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCT 730

Db 399553 TATTACCGACAAAAAAATCTCAAACATTCAAGAAATTTTACCGCTCCTTGAACAAATCTT 399612

Qy     731 TAAAACCAACCGTCCATTACTCATTATTGCAGATGATGTGGATGGTGAAGCACTTCCAAC 790

Db 399613 GAAAACAAATCGTCCACTTCTTATTGTAGCTGATGATGTTGATGGAGAAGCATTGCCAAC 399672

Qy     791 CCTTGTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATT 850

Db 399673 GCTTGTTCTTAATAAGATTAAAGGTGTCTTCAATGTTGTAGCTGTAAAAGCACCAGGATT 399732
```

```
Qy     851 TGGTGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGAT 910
           ||||||||||||||||||||    | |||||  | ||||||||||||||||||||||||| ||
Db  399733 TGGTGATCGTCGTAAAGCTCAATTGGAAGATTTGGCTATCTTGACAGGTGGTACAGTAAT 399792

Qy     911 TACAGAGGATCTAGGACTTGAATTAAAAGATGCTACAATGACAGCCCTTGGACAGGCTGC 970
           ||| || || || || |||||  | ||||||||||  |   |||  | ||||| || ||
Db  399793 TACTGAAGAACTTGGTCTTGATCTTAAAGATGCTACTCTTGAAGCTTTAGGACAAGCAGC 399852

Qy     971 TAAGATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGC 1030
           ||||   || || ||||||||   |||| ||||||||||||| | ||   | | || ||
Db  399853 TAAGGCAACTGTGGATAAAGACCACACAACAATTGTTGAAGGTGCTGGTTCTGCTGATGC 399912

Qy    1031 TATTGCTAACCGTATTGCACTGATTAAATCGCAATTAGAAACAACAACTTCTGACTTTGA 1090
           |||| || | ||| | ||  | |||||| ||||| | |||| ||||||||||||  |||||
Db  399913 TATTTCTGATCGTGTAGCTATTATTAAAGCGCAAATTGAAAAAACAACTTCTGATTTTGA 399972

Qy    1091 CCGTGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGT 1150
           |||||||| |||||||||| | || ||||||||||||||||| || ||| | |||||
Db  399973 TCGTGAAAAATTACAAGAACGTCTTGCAAAATTAGCTGGTGGTGTTGCAGTTGTTAAAGT 400032

Qy    1151 AGGAGCTCCAACAGAGACAGCTTTAAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAA 1210
           || ||| |||| || ||||   | |||| |||||||||    ||||| ||||| || ||
Db  400033 TGGTGCTGCAACTGAAACAGAGCTCAAAGCAATGAAACTCTTGATTGAAGATGCACTTAA 400092

Qy    1211 TGCTACACGTGCAGCCGTTGAAGAAGGTATCGTTGCTGGTGGTGGAACAGCACTTATTAC 1270
           ||| ||||||||||||||||| |||||||| ||| |||||||||| |||||||| |||
Db  400093 TGCAACACGTGCAGCCGTTGAGGAAGGTATTGTTTCTGGTGGTGGTACAGCACTTGTTAA 400152

Qy    1271 GGTTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGGACGTAACAT 1330
           | |||||      | |     |||      || || |||  |   || ||   ||||||
Db  400153 TGCTATTGCTGCTTTGGATAAACTTTCAGAAGAAGGAGATATTCAAACAGGTATTAACAT 400212

Qy    1331 TGTGCTTCGTGCTCTAGAAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGG 1390
           ||| |  |||||||| ||||| || || |||||||||||   ||||| || || |||||
Db  400213 TGTTCGCCGTGCTCTTGAAGAACCAGTTCGTCAAATTGCTGCCAATGCAGGATATGAAGG 400272

Qy    1391 CTCCGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAAC 1450
           || ||  | ||||| ||  |             | ||| |||||||| ||||||||||
Db  400273 TTCAGTTATCATTGATAAACTTCGTTCAGAAGAAGTAGGTACAGGATTCAATGCTGCAAC 400332

Qy    1451 AGGTGAGTGGGTTGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAACACGATC 1510
           ||| | |||||| |||||||| || |||||| |||||||||  |||||||| || ||
Db  400333 TGGTCAATGGGTTAATATGATTGAAGAAGGAATTGTTGACCCTGCAAAAGTAACTCGTTC 400392

Qy    1511 AGCGCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAAGCAGTTGTTGC 1570
           |||  | ||||| || |||||||| ||| || | |||||||| || |||||||| |||||
Db  400393 AGCTTTGCAAAACGCGGCTTCTGTTGCTGGTTTAATTTTGACTACTGAAGCAGTAGTTGC 400452

Qy    1571 TAATAAACCTGAACCAGCTACGCCAGC 1597
           |||||| || |||||||||| | |||||
Db  400453 TAATAAGCCAGAACCAGCTGCTCCAGC 400479


Search completed: April 15, 2003, 01:18:55
Job time : 3259.59 secs
```